Incidence and Subtypes of Human Metapneumovirus among Selected Rural and Urban Populations of Kenya, 2006 - 2009

Victor Oduor Omballa

A Thesis Submitted in Partial Fulfillment for the Degree of Master of Science in Laboratory Management and Epidemiology in the Jomo Kenyatta University of Agriculture and Technology

2012

DECLARATION

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

Signature …….. **Date ……**

Victor Oduor Omballa

This thesis has been submitted for examination with our approval as university supervisors

Signature …….. **Date ……**

Kariuki M. Njenga

CDC, Kenya

Signature …….. **Date ……**

Juliette R. Ongus

JKUAT, Kenya

Signature …….. **Date ……**

Joseph Oundo

CDC, Kenya

DEDICATION

This work is dedicated to the families and specifically children affected by respiratory illness associated with viral etiologies in Kenya.

To my mother, Emma Akello, grandfather, the Late James Oduor, uncles, aunties and all my cousins for their unwavering support that enabled me to complete this work.

ACKNOWLEDGEMENTS

Firstly, I would like to give special thanks to Lord God Almighty for enabling me to come this far and giving me the ability, knowledge and wisdom. I wish to express my deepest gratitude to my supervisors, Dr. Kariuki M. Njenga, Dr. Joseph Oundo, and Dr. Juliette R. Ongus for their guidance and input throughout the period of this project.

Secondly, I would wish to thank all teaching and non-teaching staff in the Field Epidemiology and Laboratory Management Training Program (FELTP) – Kenya and JKUAT. They provided support to see this project through. Dr. Breiman Rob, Dr. Feikin Danny, Dr. Omolo Jared, Otieno Cosmas Leonard, Audi Allan, Aura Barrack, Dr. Bulimo Wallace and Dr. Waiboci-Muhia Lilian many thanks for the support.

I am indebted to the Centers for Disease Control and Prevention, Kenya and the Kenya Medical Research Institute, Nairobi for making the research possible through financial and material support.

I appreciate the support of Gilbert Kikwai and Sam Symekher, who assisted in virus isolation. Special thanks also to the excellent GDD/IEIP laboratory team; Gikundi Solomon, Nderitu Leonard, Gikunju Stella, Wamburu Kabura, Odhiambo Dennis, Wachira Cyrus, Wakhule Lilian, Mayieka Lilian, Kang'ara Agnes, Shigoli Triza, Wambui Beatrice, Nindo Fredrick, Abdullahi Hashim, Ogolla Melvin, Kosgey Abigail, Ahenda Petronilla, Leitich Imelda, Onono Francis for

their support when the going got tough, the fun, the ups and downs, and above all for teamwork.

Many thanks to the Global Aids Program – Kenya laboratory team; Makhoha Ernest, Baraza Moses, Akinyi Nelly, Chege Judy, Kanyara Lucy, Nyaga Pauline, Odawo Linus and Ochieng Edwin for their advice.

I may not have made it without the tremendous support of friends; Ogada Steve, Opiyo Veronica, Mwaisaka Elisha, Ndanu Brenda, Moraa Catherine, Sempele Noelle, Gangla Dorry and Dr. Odongo David who were an inspiration in more ways than one. To all of them I say thank you and God bless you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND ACRONYMS

ARTI	Acute respiratory tract infections
aMPV	Avian metapneumovirus
BHQ 1	Black hole quencher dye 1
CAP	Contig assembly program
CDC	Centers for Disease Control and Prevention
CO_2	Carbon dioxide
CPE	Cytopathic effect
C_T	Crossover threshold
dNTP	Deoxy nucleotide triphosphate
ddNTP	Dideoxy nucleotide triphosphate
DMEM	Dulbecco's minimum essential media
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FAM	6-carboxy-fluorescein
FBS	Fetal bovine serum
GTR	General time reversible
hMPV	Human metapneumovirus

hRSV Human respiratory synctial virus

KEMRI Kenya Medical Research Institute

LLC-MK2 Rhesus monkey epithelial kidney cells

LRTI Lower respiratory tract infections

mL milli Litre

mM milli Molar

NCBI National Center for Biotechnology Information

NP Nasopharyngeal swab

OP Oropharyngeal swab

PCR Polymerase chain reaction

RPM Revolutions per minute

RR Risk Ratio

rt real time

RNA Ribonucleic acid

RT Reverse transcription

SAS Statistical analysis system

Taq *Thermus Aquaticus*

URTI Upper respiratory tract infections

µM	micro Molar
µL	micro Litre

ABSTRACT

Human metapneumovirus (hMPV) discovered recently is a member of the family *Paramyxoviridae* viruses which are responsible for acute respiratory tract infections (ARTI) in young children, elderly patients, and immuno-compromised hosts. Studies conducted in various populations have suggested a prevalence of hMPV infection estimated at 5-15%. Epidemiological data and genetic diversity on the virus are documented elsewhere, but not in developing countries especially Africa. In developing countries, the public health importance of viral etiologies is underestimated and the disease burden is not well documented. This study investigated the incidence rates of hMPV infections among two selected populations in Kenya; Lwak, a rural community and Kibera, an urban informal settlement, each with an estimated population of 27,000. Nasopharyngeal and oropharyngeal swabs were collected from patients with acute respiratory tract infections in these study sites and screened for hMPV by real time reverse transcription polymerase chain reaction. The study identified the most affected age as children less than 5 years old with incidence rates of 7.59 per 1000 person years in males below 12 months in Kibera. Overall, incidence rates per 1000 person-years for hMPV in the different study sites were 1.12 for Kibera and 0.57 for Lwak. Using the young adults (18 – 34 years old) as reference; children 12 – 23 months from Lwak were 12 times (R.R = 11.9, p-value < 0.001) more likely to be infected by hMPV.

Children < 12 months from Kibera were 18 times (R.R = 18.0, p-value < 0.001) more likely to be infected by hMPV. Adjusted incidence rates by percentage of cases with acute respiratory tract infection (ARTI) whose sample were taken and those who visit the field clinic indicate that children < 12 months from Lwak had the highest adjusted incidence rate of 29.5 per 1000 person-years of observation. Subtypes of hMPV circulating in the two study populations were also characterized. Of the four subtypes of hMPV, the common subtype of hMPV circulating was B2 (10/17, 59%), followed by A2 (5/17, 29%) and B1 (2/17, 12%). Subtype A1 was not detected among the 17 hMPV isolates. The common clinical symptoms observed among hMPV-infected cases included cough, fever, runny nose, sneezing, temperature ≥ 38˚C and Oxygen Saturation ≤ 90. In conclusion, hMPV is present in the selected Kenyan populations, contributing to morbidity as high as 0.86 per 1000 persons in all ages at both the selected urban and rural settings. Most of the major subtypes of hMPV are circulating in the selected study populations.

CHAPTER ONE

1.0. INTRODUCTION

The human metapneumovirus (hMPV) is a member of the *Paramyxoviridae* viruses which are associated with acute respiratory tract infections (ARTI) in young children, elderly patients, immuno-compromised and lung transplant individuals (Boivin *et al*., 2004; Dare *et al.*, 2007). The hMPV is a major cause of human morbidity and mortality worldwide, causing a spectrum of illnesses ranging from upper respiratory tract infections (URTI), lower respiratory tract infections (LRTI) to severe bronchiolitis and pneumonia. The virus is common in children less than 5 years of age but also found in adults of all age groups (Falsey *et al*., 2003; Principi *et al*., 2006). More than 90% of children show serological evidence of hMPV infection by the time they are 5 years old (van den Hoogen *et al.*, 2001; Peret *et al.*, 2002; Kahn, 2006). The virus was first identified in the winter of 2001 in the Netherlands among children suffering from respiratory tract infections (van den Hoogen *et al.*, 2001). The paramyxovirus was classified into *Pneumoviridae* sub-family and *Metapneumovirus* genus, based on gene arrangements and phylogenetic analysis (van den Hoogen *et al*., 2001).

The hMPV has a seasonal distribution with disease peaks during the cold winter periods in the northern hemisphere and fewer cases in the summer period (Peret *et al.*, 2002; Kaida *et al.*, 2006; Barenfanger *et al.*, 2008; Gaunt *et al.*, 2009).

The virus has consistently been detected in subsequent winter periods at rates between 2 – 10% in adult populations only (Walsh *et al.,* 2008).

Outbreaks usually in children associated with hMPV are localized and clinical symptoms can easily be confused with Human respiratory syncytial virus infection (Gerna *et al*., 2005; Agapov *et al.,* 2006; Kaida *et al.,* 2006). Although the clinical manifestations of respiratory tract infections are easily recognized worldwide, the etiological agent responsible for the disease is often not identified. Overall, there are two hMPV groups isolated from various regions of the world i.e. A and B, which are genetically similar, with less than 17 % nucleotide sequence differences between them (Huck *et al.,* 2006). Within the two groups, there are four subtypes of hMPV, namely A1, A2, B1 and B2 (Boivin *et al.,* 2004). The nucleotide sequence difference between these subtypes is less than 20 % whereas the average predicted amino acid sequence difference is less than 10% (Biachessi *et al.,* 2003). Recently, subtype A2 has further been classified as A2a and A2b (Huck *et al.,* 2006). A population-based surveillance was started by the Kenya Medical Research Institute and Centers for Disease Control and Prevention (KEMRI/CDC) in 2004 to determine the etiologies of respiratory tract infections in one selected urban and another rural population of Kenya. Clinical evaluations were carried out within the population to determine patients for referrals to the nearest health care facilities.

Patients meeting the inclusion criteria provide nasopharyngeal and oropharyngeal samples for testing and were offered free treatment and referrals to better hospitals in the event of an admission. Since information on the incidence of hMPV and other respiratory viruses circulating in the country were not available, this study was undertaken to determine the incidence rates of hMPV in the selected populations and to characterize the subtypes of the virus circulating in the country.

1.1. Problem Statement

Acute respiratory tract infections continue to be the leading cause of acute illnesses worldwide and remain the most important cause of infant and young children morbidity and mortality, accounting for up to 4 million deaths annually especially in developing countries (Girard *et al.,* 2005). Of this, pneumonia which is partly caused by hMPV, accounts for 1.9 million deaths annually (Scott *et al.,* 2008).

The common etiologies for pneumonia are bacterial (i.e. *Streptococcus pneumoniae* and *Haemophilus influenzae* type B) and viral infections but mortality attributed to viral etiologies has not been well elucidated (Lin *et al.,* 2005; Scott *et al.,* 2008).

In western Kenya, among rural hospitalized children, up to 282/2466 deaths (11.4%) from pneumonia were reported in a period of 2 years. These pneumonia cases peaked with two malaria seasons (Tornheim *et al.,* 2007). For many bacterial infections there are safe and efficient vaccines developed as opposed to viruses (Girard *et al.,* 2005). The most common viral etiology for ARTI is reported as human respiratory syncytial virus (hRSV) closely followed by hMPV (Viazov *et al.,* 2003).

These infections are normally self-limiting in adults, but the infections may be associated with significant morbidity in infants, immuno-compromised individuals and those with underlying conditions such as asthma, prematurity and cardiopulmonary disease. Most cases of ARTI occur in the first year of life and require hospitalization and mechanical ventilation (Williams *et al.,* 2004). Studies indicate that hMPV is an important cause of respiratory infections; however its prevalence, incidence and seasonality in Kenya have not been determined. Limited data indicates that both hMPV groups (i.e. A and B) can circulate in a single season with the possibility of a predominant subtype switching in successive seasons and outbreaks of the virus are normally localized in different communities (Kahn, 2006).

1.2. Justification for the study

The greatest burden of diseases occurs in the developing countries (Black *et al.,* 2003). Pneumonia is documented as the leading cause of death in Kenyan children (Mirza *et al.,* 1990). Information is lacking on the incidence and etiologies of important infectious diseases in developing countries despite their substantial burden. Most information about infectious disease has previously been in relation to hospital-based studies or passive surveillance systems which are limited by health-seeking habits of a population. The hMPV causes acute respiratory tract infections among various populations worldwide (Bastien *et al.,* 2004; Ludewick *et al.,* 2005; Chung *et al.,* 2006; Kahn, 2006; Mackay *et al.,* 2006; Hopkins *et al.,* 2008).

Since its discovery, the complete epidemiology and clinical features of hMPV infection and impact on health have not been fully described in developing countries especially Sub-Saharan Africa. This study was to highlight the importance of the hMPV in the selected population. Preliminary results by KEMRI/CDC population-based surveillance indicate its presence in Kenya.

The population based surveillance was initiated to determine the burden of disease of clinically defined pneumonia, diarrhea, jaundice and fever. Incidence, genetic diversity and circulating subtypes of the virus have previously not yet been documented in Kenya and this can guide in health care policy making and provide further information on pneumonia viral etiologies. This study will provide subtype information which can assist public health officials in deciding the composition of an appropriate vaccine when one becomes commercially available and also preliminary data in relation to incidence rates and clinical symptoms associated with the virus within the selected populations of Kenya.

1.3. Objectives

1.3.1. General Objective

To determine the incidence rate of hMPV in selected Kenyan populations and identify the circulating subtypes in the selected populations.

1.3.2. Specific Objectives

1. To determine the incidence rate of hMPV among selected rural and urban populations of Kenya.
2. To identify the clinical symptoms associated with hMPV infection among selected rural and urban populations of Kenya.
3. To isolate and determine the genetic subtype of the hMPV circulating strains in the selected populations.

CHAPTER TWO

2.0. LITERATURE REVIEW

2.1. Prevalence and incidence of hMPV

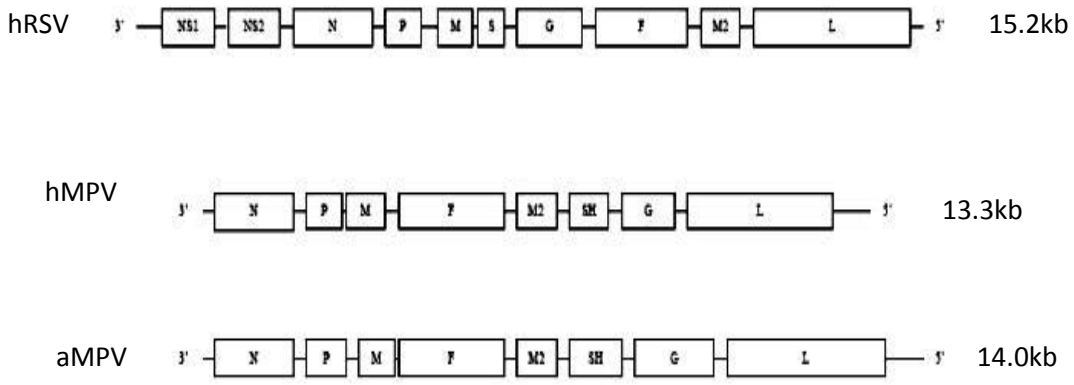
The hMPV appears to be ubiquitous across the world and it is estimated that approximately 70% of children are infected with hMPV by the time they are 5 years old (van den Hoogen *et al.,* 2001; Ludewick *et al.,* 2005). Respiratory infections are the most common illnesses among children less than 5 years throughout the world with viral etiologies primarily associated with hMPV, hRSV, Influenza virus, and Adenovirus (Murray *et al.,* 2001; Madhi *et al.,* 2007). Most studies indicate a hMPV prevalence of 5 – 15% among children less than 5 years, but morbidity attributed to hMPV is not well defined (Bastien *et al.,* 2003; Esper *et al.,* 2004; McAdam *et al.,* 2004). The hMPV is reported to be the second most frequent virus after hRSV causing ARTI in infants < 2 years (Viazov *et al.,* 2003; Chano *et al.,* 2005; Garcia *et al.,* 2006). However, an incidence of hMPV as high as 7.6/1000 person in children below 2 years old has been reported (Heikkinen *et al.,* 2008).

2.2. Avian and Human metapneumoviruses

Avian metapneumovirus (aMPV) closely resembles hMPV in terms of genomic organization and both viruses are classified in the *Metapneumovirus* genus and Paramyxoviridae family because of limited sequence data (van den Hoogen *et al.*, 2001). Both aMPV and hMPV are subdivided into four subgroups each; A, B, C and D for aMPV and A1, A2, B1 and B2 for hMPV (van den Hoogen *et al.,* 2001; Alvarez *et al.,* 2003). The aMPV was first isolated from turkeys in South Africa in 1978 while hMPV was isolated from children in Netherlands in 2001 and later these were found to have a global distribution (Buys *et al.,* 1980; van den Hoogen *et al.*, 2001; Lwamba *et al.*, 2005). Both viruses tend to have a seasonal distribution (Jones *et al.*, 1996; Kahn, 2006). Clinical symptom comparison reveals that aMPV is characterized by acute upper respiratory tract infection, cough and nasal discharge with a mortality of up to 30% while hMPV causes both lower and upper respiratory tract infections including fever, cough, runny nose and chest pain and finally developing to pneumonia (Broor and Bharaj, 2007). Full genome analysis of aMPV C indicates that it is the closest in genome similarity to hMPV compared to the other aMPV subtypes (van den Hoogen *et al.,* 2001; Alvarez *et al.,* 2003; Njenga *et al.,* 2003; Lwamba *et al.,* 2005). The overall amino acid sequence identity between aMPV and hMPV subtypes ranges from 56% to 88% for open reading frames N, P, M, F, M2-1, M2-2, and L (van den Hoogen *et al.,* 2002).

2.3. Genomic structure of hMPV

The hMPV genome consists of 8 genes in the order of nucleocapsid (N), phosphoprotein (P), Matrix (M), Fusion (F), Second Matrix (M2), Small Hydrophobic (SH), attachment protein (G) and the Large Polymerase (L) similar to aMPV (van den Hoogen *et al.*, 2001). The length of the entire genome of hMPV and aMPV are approximately 13.3 and 14.0 kilobases (kb) respectively. The aMPV has the SH, G and L genes slightly larger than the same genes in the hMPV genome (Lwamba *et al.*, 2005). The genomic arrangement of both viruses is similar to hRSV (Biachessi *et al*., 2003; Dare *et al.*, 2007). Compared to the genomes of hRSV, the hMPV and aMPV genome lack the two non-structural proteins, NS1 and NS2 at the 3'-end and the hMPV and aMPV have a different gene order as shown in Figure 2.1 (van den Hoogen *et al.*, 2002). The hRSV genome is approximately 15.2kb (Figure 1.0). The major difference between group A and B of hMPV occurs due to nucleotide polymorphisms at the G and SH genes (Biacchesi *et al.*, 2003). These genes are the least conserved containing only a 32-37% amino acid identity of the G protein between groups A and B (Agapov *et al.*, 2006).



(van den Hoogen *et al.,* 2002)

Fig 2.1: Genomic arrangement of hMPV, aMPV and hRSV

The hMPV is genetically similar to aMPV; a turkey virus classified in the same genus as hMPV.

2.4. Classification/subtypes

Sequencing and phylogenetic analysis indicated the presence of two major groups of hMPV designated as A and B (Esper *et al.,* 2004; Chano *et al.,* 2005) based on the M, F, G, N and L genes. The major difference in the two groups is due to nucleotide polymorphisms in the G and SH genes. The G-gene has sequence variability among groups and subtypes. Subsequently, four subtypes i.e. A1, A2, B1 and B2 were described and full length sequences of the genomes reported (van den Hoogen *et al.,* 2002; Biacchesi *et al.,* 2003).

2.5. Seroprevalence of hMPV

Infection by hMPV is common in children < 5 years old (Kahn, 2006). Seroprevalence studies of hMPV showed that the virus had been in circulation for approximately 50 years (van den Hoogen *et al.,* 2001). By the age of 5 years, more than 90% of children test positive for hMPV and by adulthood nearly 100% are serologically positive for hMPV (van den Hoogen *et al.,* 2001; Kahn, 2006). Infants < 3 months old have hMPV antibodies that are maternally derived (Kahn, 2006). However, it has not been determined whether this confers immunity or reduces pathogenesis.

2.6. Risk factors associated with hMPV

Risk factors to respiratory illnesses have previously been defined in developed nations including parental smoking during pregnancy, presence of pets, premature child births, chronic lung disease and congenital heart disease which have not been described in developing nations (McConnochie *et al.,* 1988). In the developed countries, improvements in living conditions and air quality have significantly reduced the occurrence of ARTIs. A hospital-based study in developing countries described the risk factors of ARTI to include; age ≤ 3 months, indoor air pollution, recurrent ARTI infection and using cooking fuels other than gas (Al-Sonboli *et al.,* 2006).

It has also been established that zinc supplementation significantly reduces the incidence of pneumonia in children less than 12 months old (Brooks *et al.,* 2005). Lack of zinc in the diet can be a risk factor in acquiring pneumonia from viral etiologies.

2.7. Other medical complications and hMPV

The hMPV has been associated with pathogenesis of acute otitis media with 61% on hMPV-infected children younger than 3 years old developing otitis media (Williams *et al.,* 2006; Heikkinen *et al.,* 2008). The virus also predisposes children to bacterial infections and is associated with encephalitis and encephalopathy after isolation from brain and lung tissues of a severe case (Lin *et al.,* 2005; Schildgen *et al.,* 2005).

The hMPV has been found in a case report of an immuno-compromised child causing prolonged and severe respiratory complications in subsequent consecutive seasons leading to death (Pelletier *et al.,* 2002). Studies have indicated an exacerbation of asthma and chronic obstructive pulmonary disease (COPD) in relation to hMPV (Williams *et al.,* 2004). The virus is implicated in severe lower respiratory disease in organ transplant patients (Dare *et al.,* 2007).

2.8. Vaccine development

A safe and effective vaccine against hMPV is under development (Tang *et al.*, 2003). A bivalent live recombinant vaccine of human parainfluenza virus-3 with F-gene of hMPV has been shown to induce hMPV-specific antibodies in hamsters (Tang *et al.*, 2003; Skiadopoulos *et al.*, 2004). The successful development of a vaccine candidate will significantly reduce the burden of disease from the two viruses.

CHAPTER THREE

3.0. MATERIALS AND METHODS

3.1. Study Design

This was a descriptive study to determine the incidence rate of hMPV and identify the circulating subtypes of the virus.

3.2. Study Sites

The patients were from two study populations which were Lwak rural community and Kibera urban informal settlement (Fig 3.1).

Fig 3.1: Map of Kenya indicating Lwak and Kibera study sites

3.2.1. Lwak Rural Community

Lwak approximately at latitude 0.14°S and longitude 34.36°E is a rural community located in Rarieda District of Nyanza province (Fig 3.2). The KEMRI/CDC population-based surveillance serves an estimated population of 25,000 enrolled participants from 33 villages living within approximately 5 kilometers radius of the Lwak mission hospital, the designated health care facility. Most Kenyan population seek medical care within 5 km of their residence (Noor *et al.*, 2003).

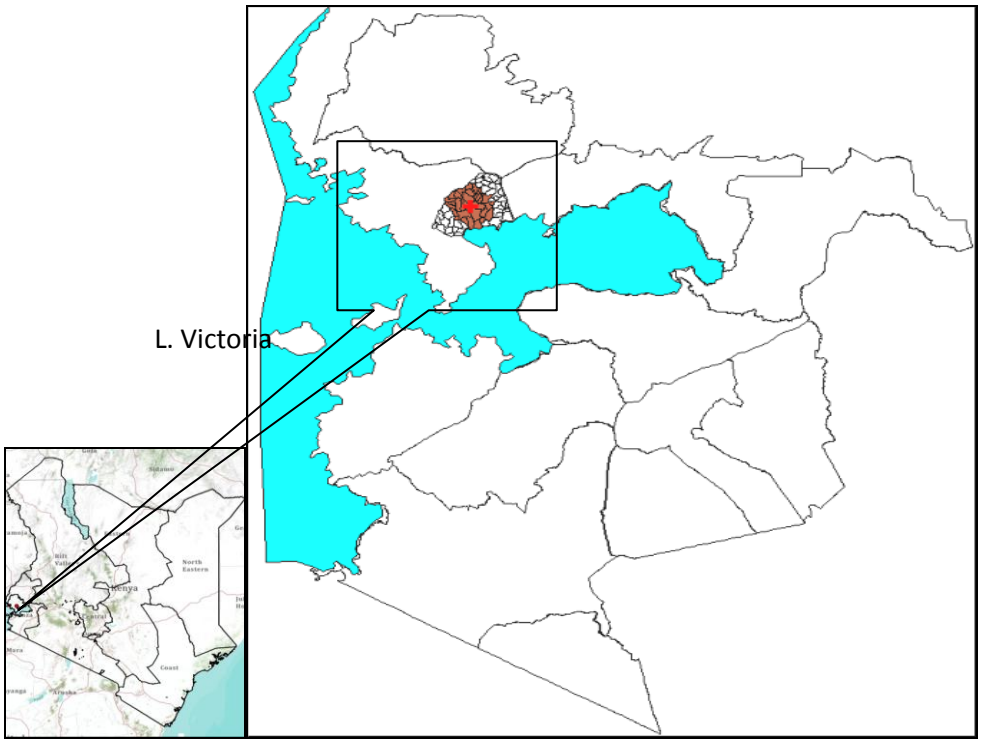


Fig 3.2: Map of Nyanza province of Kenya indicating Lwak study site

3.2.2. Kibera Urban Informal Settlement

Kibera approximately at latitude 1.32°S and longitude 36.78°E is an urban informal settlement located in Nairobi province of Kenya. The slum is located roughly 5 kilometers southwest of the city centre of Nairobi (Fig 3.3). The study site encompassed an area of 2.5 square kilometers accounting for less than 1% of Nairobi's total area, but holds more than a quarter of its population. The KEMRI/CDC population-based surveillance only covers 2 out of 13 major villages that include Soweto and Gatwekera lying southwest of Kibera with a total population of approximately 27,000 enrolled participants. Kibera was selected due to a relatively variable ethnic community, lack of adequate social amenities with a less mobile population as compared to other villages. Other villages not under surveillance include Kianda, Kisumu Ndogo, Lindi, Laini Saba, Siranga/Undugu, Makina and Mashimoni. Majority of these populations are persons from the rural community seeking employment.

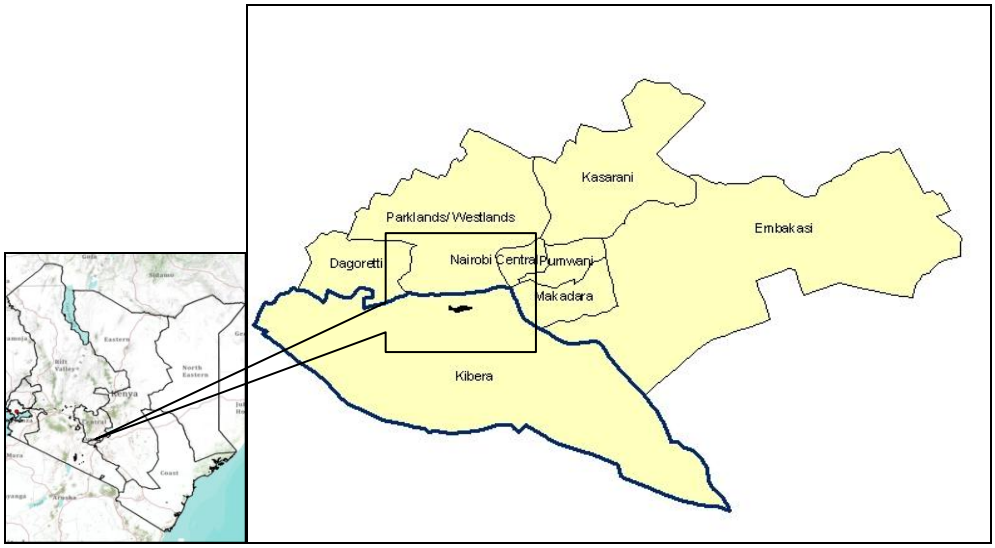


Fig 3.3: Map of Nairobi province of Kenya indicating Kibera study sites

3.3. Study Population

This was a population-based surveillance conducted at 2 sites; two villages of Kibera urban informal settlement consisting of a population of approximately of 27,000 and Lwak rural community; 33 villages with a population estimated at 25,000 persons. The populations were previously determined by the KEMRI/CDC population-based surveillance project by enumeration and recruiting of consenting individuals into the population under surveillance (Feikin *et al.*, 2011). Population-based surveillance normally provides accurate information on burden of disease and incidence rates. Calculation of incidence rates can be compared in different locations as a measure of disease burden. The total population under the study were previously enumerated and registered as members of the population-based surveillance prior to the study.

3.4. Inclusion criteria

Any patient enrolled in the KEMRI/CDC population-based surveillance between 1st October 2006 and 30th September 2009 presenting at any of the two clinics with ARTI symptoms. Nasopharyngeal and oropharyngeal (NP/OP) swab samples were only collected from symptomatic patients. For children < 5 years, these included cough, difficulty in breathing, inability to drink, vomiting, convulsions, lethargy, in-drawing, stridor, oxygen saturation ≤ 90, temperature ≥ 38˚C and sore throat. In children > 5 years and adults, the symptoms included any patient presenting with cough, difficulty in breathing, sore throat, temperature ≥ 38˚C and oxygen saturation ≤ 90.

Exclusion Criteria – Any person within the KEMRI/CDC population-based surveillance not presenting with respiratory illness at the health care facility and those that are outside the population-based surveillance.

3.5. Experimental procedures

3.5.1. Case finding

Biweekly household visits were carried out in both Lwak rural community and two villages of Kibera urban informal settlement in Nairobi. The suspected cases were referred to free study clinics within the surveillance area: Lwak mission hospital for the rural population and Tabitha medical clinic for Kibera urban population.

Sick people within the two populations could also voluntarily visit the clinics, which provided free treatment. Patients presenting at the clinics were easily identified with their enrollment identity. A questionnaire was administered to a suspected case following consent (Appendix B). The questionnaire captured information on patient details, demographic location, vital signs and clinical information.

3.5.2. Specimen collection

Both nasopharyngeal and oropharyngeal swabs were collected from each respondent meeting the case definition and then these were placed in the same cryovial containing 1.0 mL viral transport medium (VTM) and transported through a cold chain to the KEMRI/CDC Laboratory in Nairobi, Kenya for testing. After initial testing, all specimens were kept at -80˚ C freezer for possible future testing. Archived samples were reviewed retrospectively before 1st January 2009 and thereafter, other samples were collected and analyzed progressively.

3.5.3. Screening ARTI patients for hMPV

Nucleic acid was extracted from the NP/OP swabs collected from ARTI cases using the QIAamp RNA Extraction kit (Qiagen Inc, Amsterdam, Netherlands) according to the manufacturer's protocol (Appendix H).

The extracted RNA was tested for hMPV and other respiratory viruses in a one-step real time RT-PCR system using AgPath-ID One-step RT-PCR kit (Applied Biosystems, Foster City, CA) according to the manufacturer's protocol.

A 97 base pair fragment of the F-gene was amplified. The F-gene coding for the fusion protein is integral in viral attachment and is a unique identifier of the paramyxvoviruses (van den Hoogen *et al.*, 2002; Schowalter *et al.*, 2006). It was detected using a fluorescent-labeled hydrolysis/Taqman probe at a final concentration of 0.1µM (Table 3.1). The probe was flanked on either end by forward and reverse primers at a concentration of 0.8µM each for amplification by AmpliTaq® DNA polymerase Gold. Prior to amplification, cDNA was formed by a reverse transcription enzyme mix using 5 µL of previously extracted RNA in a 25µL RT-PCR reaction mix. The reaction mix was optimized to include 1X RT-PCR buffering system. Positive and negative controls were always run with the samples to assess nucleic acid amplification. Amplification conditions included reverse transcription at 45°C for 10 min, AmpliTaq® Polymerase activation at 95°C for 10 min, then 45 cycles of denaturation at 95°C for 15 sec and primer annealing/extension at 55°C for 1 min. Lower C_T values indicate strong positives as opposed to higher C_T values. Any sample that did not test positive for hMPV by real time RT-PCR did not undergo further testing.

Table 3.1: Primers and probes used for identification of hMPV by Conventional and real time RT-PCR

CONVENTIONAL	**OLIGONUCLEOTIDE PRIMER**
HMPV-F	5’ GAG CAA ATT GAA AAT CCC AGA CA 3’
HMPV-R	5’ GAA AAC TGC CGC ACA ACA TTT AG 3’
REAL TIME	**OLIGONUCLEOTIDE PRIMER/PROBE**
HMPV-F	5’ CAA GTG TGA CAT TGC TGA YCT RAA 3’
HMPV-R	5’ ACT GCC GCA CAA CAT TTA GRA A 3’
HMPV – P	5’ TG GCY GTY AGC TTC AGT CAA TTC AAC AGA 3’

Key: F = Forward primer

R = Reverse primer

P = Fluorescent probe

The conventional primers (Falsey *et al.*, 2003) were used at a final concentration of 0.3µM targeting the highly conserved F-gene coding for the Fusion protein in a 50µL reaction mix. The real time RT-PCR primers and probes for hMPV screening were designed to target a fragment of the F-gene (3549 – 3628) that codes for the Fusion protein with an expected product size of 97bp amplifying all the subtypes of hMPV.

The probe was labeled with **FAM** dye on the 3’-end and **BHQ1** on the 5’-end.

3.5.4.Virus Isolation

Samples that were hMPV positive were subjected to viral isolation in confluent rhesus monkey epithelial kidney cells (LLC-MK2) to increase viral titers. The infections were carried out using 300µL of specimen pre-diluted with maintenance media (Appendix C) in a ratio of 1:2 after washing in trypsinized wash buffer three times. These were incubated at 35˚C for 1h for adsorption of viral particles to the cell monolayer and then 5mL of Dulbecco's Minimum Essential Media (DMEM) which had previously been supplemented with 10% Fetal Bovine Serum (FBS), 1% L-Glutamine and 1% antibiotics was added to each flask. Flasks were then incubated at 35°C in a 5% CO_2 with a humidity of 85 % and monitored for cytopathic effect (CPE) for up to 21 days. Cells showing CPE were harvested and centrifugation at 3000 revolutions per min (RPM) for 10 min and the supernatant stored at -80°C prior to viral RNA extraction and sequencing.

3.5.5. Conventional RT-PCR

In order to obtain a hMPV DNA fragment for sequencing, conventional RT-PCR was carried out using Superscript III One-step RT-PCR system with Platinum *Taq* High Fidelity, according to manufacturer's instructions. The RT-PCR system consists of Superscript III RT/Platinum *Taq* High Fidelity enzyme mix for both cDNA formation and PCR amplification.

The 2X reaction mix was optimized by 2.4 mM of Mg^{2+}, 0.4 mM of each dNTPs and stabilizers. RT-PCR was carried out in a 9700 ABI thermocycler (Applied Biosystems) as follows; Reverse Transcription at 48°C for 30 min, *Taq* Polymerase activation at 95°C for 10 min, then 45 cycles of denaturation at 94°C for 15 sec, primer annealing at 55°C for 1 min and extension at 72°C for 2 min. A final extension was carried out at 72°C for 7 min and products stored at 4˚C. The hMPV -positive and -negative controls containing standardized viral RNA extract and nuclease-free water, respectively, were included in each assay.

RT-PCR products were verified by agarose gel electrophoresis. The gel was immersed in 1X TAE buffer. The products previously stained with bromophenol blue loading dye were loaded into the wells and allowed to run for 1h at 100 volts from the negative cathode to positive anode and later viewed under Ultra violet light at 312 nm wavelength.

3.5.6. Fragment recovery and purification

Amplicons were purified from the PCR reaction mix using QIAquick PCR product Extraction Kit (Qiagen, Inc, Amsterdam, Netherlands) according to the manufacturer's protocol. Briefly, the remaining 40µL RT-PCR products after gel electrophoresis were purified to remove excesses of primers, Mg^{2+}, unincorporated nucleotides and non-specific products by adding to five volumes (i.e. 200 µL) of a buffer enabling efficient binding of double-stranded PCR products (i.e. PBI buffer).

To bind the DNA, the solution was pipetted into a QIAquick column inserted in a 2.0mL collection tube and centrifuged for 1 min at 13,000rpm at Room Temperature. Later, 0.75mL of ethanol-containing PE buffer was added to the QIAquick column so as to remove all traces of excess non-PCR products and salts and centrifuged for 1 min at 13,000rpm at Room Temperature. The flow-through was discarded then the QIAquick column placed in 1.5mL nuclease-free tube and amplicon eluted with 30µl of nuclease-free water at a final centrifugation of 13,000rpm for 1 min. This was stored at temperatures below -20°C until ready for the cycle sequencing reaction.

3.5.7. DNA Sequencing

DNA sequencing was carried out in both directions using both the forward and reverse primers previously used for RT-PCR amplification. Cycle sequencing by chain termination reaction (Sanger *et al.,* 1977) was carried out prior to DNA sequencing to enable fluorescent labeling of the PCR products using dideoxynucleotides triphosphates (ddNTPs). Elongation normally occurs by addition of dNTPs, but immediately a ddNTP is added during elongation, then termination occurs. The ddNTP will later fluoresce when passed through the 3130xl Genetic analyzer camera and used to plot a sequence curve. Before loading cycle sequencing products into the 3130xl Genetic analyzer, the products were purified using Applied Biosystems Centri Sep spin columns (Princeton Separations, Foster City, CA). The Centri-sep columns are normally used for fast and efficient purification of cycle sequencing products.

Separation was carried out in a separation polymer matrix, 3130 POP 7 Performance Optimized Polymer (Framingham, MA, USA) where sequences move from cathode to anode according to length and therefore allowing the shorter fragments to move faster until sequencing stops at the fragment with the longest nucleotide sequence (Appendix D).

3.6. Data analysis

3.6.1. Epidemiologic data analysis

Epidemiologic data was analyzed by Statistical Analysis System - SAS (SAS Institute Inc., Cary, NC). For person-years calculations, residence status in the surveillance area was used to determine start and stop dates for person-time contribution. Persons that were not present during the subsequent visits for more than 2 weeks were excluded from the person-time analysis. The difference between the start and stop dates were divided by 365.25 days to generate person-years contributed to the study within the specific time period. Some individuals exit, enter or rejoin the study within the 3 year study period i.e. the population under study is dynamic.

Incidence rate for the 3 year study period was determined by the total number of hMPV cases in relation to the person-year.

Adjustments were made in relation to symptomatic patients who visited the clinics and those that NP/OP samples were collected. Descriptive statistics of age, sex and presenting clinical signs and symptoms were analyzed.

Analyses were carried out to determine which age group and sex had the highest number of positive specimens for hMPV in relation to the total number of swabs taken.

Risk Ratio was analyzed in comparison to a selected reference group to determine which age group was at a higher risk of getting hMPV infection.

3.6.2. DNA sequences data analysis

Data analysis was carried out using Applied Biosystems' data collection software v2.0 that retrieved sequences from the 3130xl machine (Appendix D) and sequence editing and contig formation was carried out by Bioedit (Hall, 1999) software. For further analysis and comparison of the generated sequences, a neighbor joining phylogenetic tree was drawn using Mr Bayes v3.1 (Ronquist and Huelsenbeck, 2003) software from 1,000,000 replicates at a sample frequency of 1000 and viewed by FigTree software v1.2.2 (Rambaut, 2006). Mr Bayes software is a program for the Bayesian inference of phylogeny using the General Time Reversible (GTR) model (Tavare *et al.,* 1986) which is considered more comprehensive and random. The phylogenetic tree was drawn to distinguish the subtypes of circulating hMPV in the selected populations in comparison to subtypes in the National Center for Biotechnology Information (NCBI) Gene Bank. Contigs were assembled using the Contig Assembly Program (CAP) of the Bioedit suite v7.0.5 (Hall, 1999).

Multiple sequence alignment was carried out using muscle v3.6 software (Edgar, 2004) before conversion to the nexus format recognized by Mr Bayes v3.1 software.

Sequence identity and divergence was analyzed using DNAstar v8.0 Megalign software (DNASTAR Inc., MadisonWI).

3.7. Data Management

3.7.1.Data entry and storage

Epidemiologic data (questionnaire) entry was carried out using personal digital assistance (PDA) to ensure data quality and storage was carried out in a Microsoft access database maintained by a data manager. Data verification was routinely carried out to identify errors in the data collection and entry and later, errors identified were reviewed appropriately. After data entry, hard copies of questionnaires were stored in restricted lockable long-term storage units at KEMRI/CDC.

3.7.2. Study Limitations

The study focused on selected study populations and not the wider Kenyan population and is therefore not representative The clinical symptoms are subject to recall bias and not all study participants visited the field clinic for collection of sample. These symptoms may also be as a result of coinfection with pathogens that were not tested for. Only 17 out of 30 hMPV isolates were recovered due to difficulties in hMPV isolation. Therefore the characterization of subtypes may not be representative of all subtypes present in Kenya. The sensitivity of the real-time RT-PCR test used to screen for hMPV in ARTI cases is 82 - 90% for different pathogens. Therefore we may have missed up to 18% of hMPV cases within the two populations.

CHAPTER FOUR

4.0. RESULTS

4.1. Demographic characteristics of the study populations

Between 1st October 2006 to 30th September 2009, the total population under study at Lwak was 27,294 (Figure 4.1), whereas that at Kibera was 27,473 (Figure 4.2) with a female population of 53% in Lwak and 50% in Kibera.

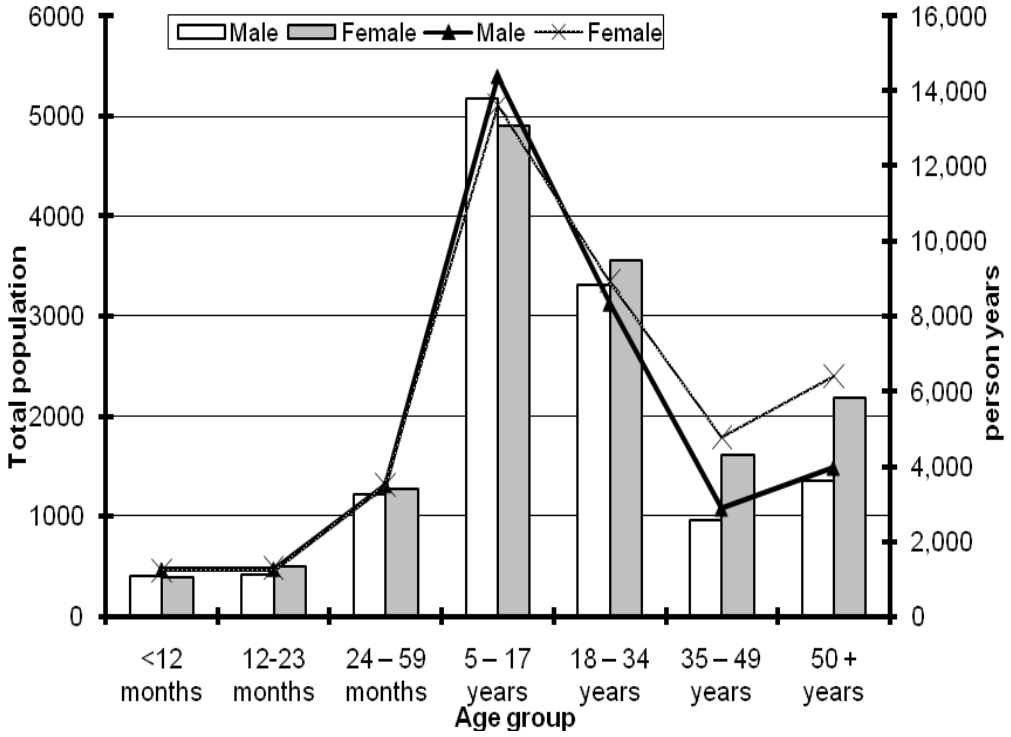


Fig 4.1: Normal distribution of the total population under study indicating person years by age group and sex in Lwak, 2006 - 2009.

Figure 4.1 indicates the total population under study (27,294) for Lwak rural community indicating a normal distribution of the population segregated by sex and age group. The bar graph indicates a male and female population that is similar with person-years represented on the right side y-axis having a similar trend. Majority (62%) of persons under the population were aged 5 – 18 years with less than 50% for the other age groups.

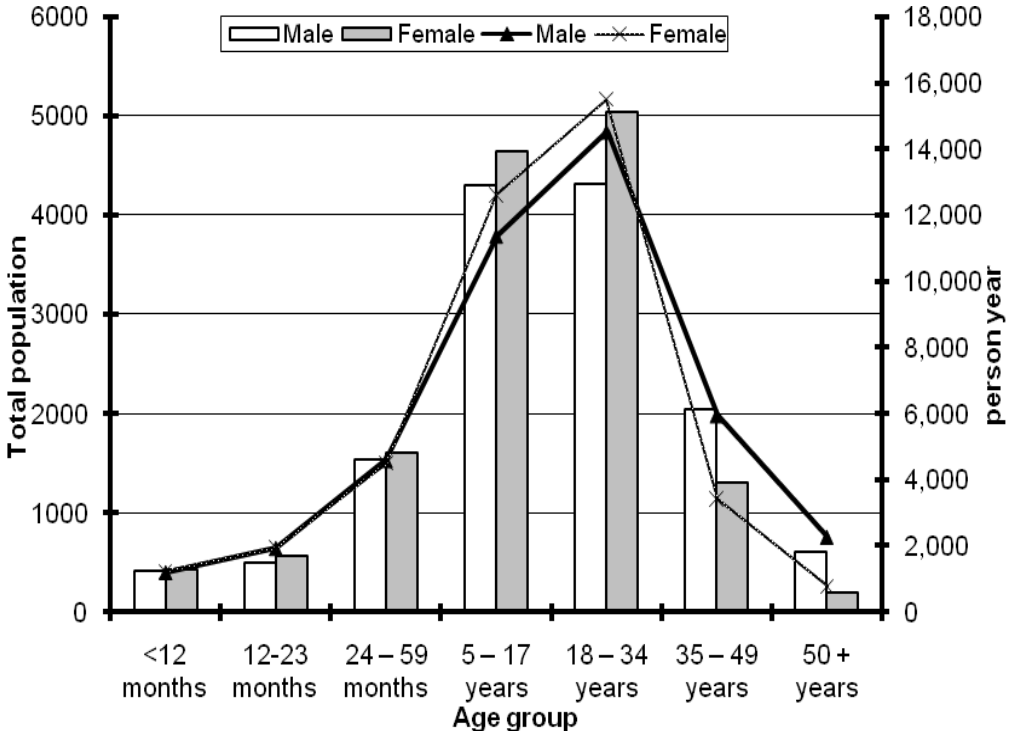


Fig 4.2: Normal distribution of the total population under study indicating person years by age group and sex in Kibera, 2006 - 2009.

Figure 4.2 indicates the total population under study (27,473) for Kibera urban informal settlement indicating a normal distribution of the population segregated by sex and age group. The bar graph indicates a male and female population that is similar with person-years represented on the right side y-axis having a similar trend. . Majority (67%) of persons under the population were aged 5 – 18 years with less than 50% for the other age groups.

4.2. Demographic characteristics of patients whose swabs were collected and positivity proportions

Between 1st October 2006 and 30th September 2009, from approximately 27,000 persons per site, a total of 1,583 and 1,794 nasopharyngeal and oropharyngeal swabs were collected from symptomatic patients attending Lwak Mission Hospital and Tabitha Health clinic in Kibera respectively (Table 4.1). A total of 1,025 and 541 samples were collected before 1st January 2009 and were archived samples. Between 1st January 2009 and 30th September 2009, 558 samples from Lwak and 1,253 samples from Kibera were collected and analyzed together with the archived samples. Specimens were collected from as young as 11 day old children with a mean age of 16 and 11 years from Lwak and Kibera respectively (Table 4.2).

Table 4.1: Demographic characteristics of patients with NP/OP swabs collected from both study sites

	NP/OP collected	Mean age in yrs	S.D * in yrs	Age range	†IQR in yrs
Lwak	1,583	16	19	11 days – 91 yrs	23
Kibera	1,794	11	13	14 days – 64 yrs	17

*S.D = Standard Deviation

†IQR = Interquartile range

Within this 3 year study period, the demographic details of the 43 hMPV positive patients from Lwak rural community and 92 hMPV positive patients from Kibera urban informal settlement were analyzed (Table 4.2). The youngest hMPV case from both study sites was 2 months old with a mean age of 11years and 7 years from Lwak and Kibera respectively.

Table 4.2: Demographic characteristics for hMPV positive patients from both study sites

	hMPV positive	**Mean age in yrs**	***S.D in yrs**	**Age range**	**†IQR in yrs**
Lwak	43	11	14	2 months – 53 yrs	10
Kibera	92	7	10	2 months – 47 yrs	6

*S.D = Standard Deviation

†IQR = Interquartile range

The total number of NP/OP swabs collected per site were analyzed for hMPV percent positive by age group and sex to determine how many of the total number of swabs collected turned positive for hMPV by real time RT-PCR (Figure 4.3 and 4.4).

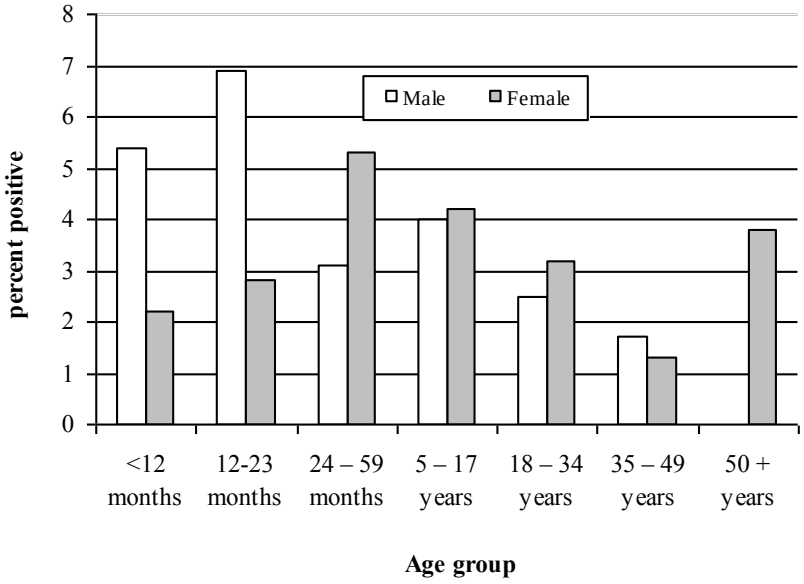


Fig 4.3: hMPV percent positive for Lwak by age group and sex

Of the total number of NP/OP swabs collected (1,583) from Lwak rural community, 602 swabs were from children aged < 5 years old and 981 swabs were taken from persons > 5 years old. NP/OP swabs were hMPV positive by real time RT-PCR across the different age groups. However, there were no swabs collected for males 50+ years old. The hMPV had the highest proportion among males aged 12 – 23 months (6.9%) and a sudden drop in males aged 24 – 59 months (3.1%) with the lowest in females 35 – 49 years (1.3%).

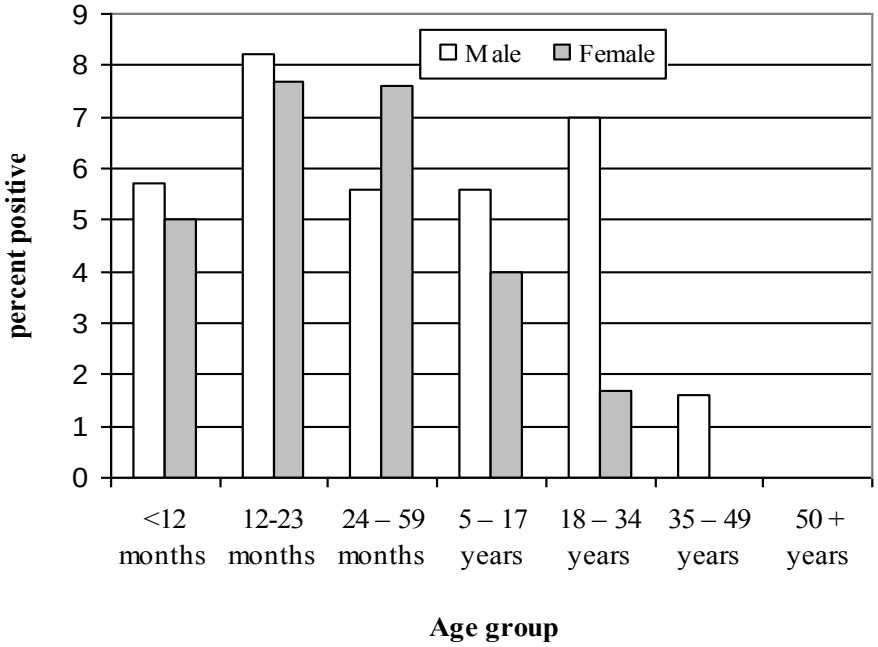


Fig 4.4: hMPV percent positive for Kibera by age group and sex

Of the total number of NP/OP swabs collected (1,794) from Kibera urban informal settlement 917 swabs were taken from children < 5 year old and 877 swabs from persons > 5 year old. NP/OP swabs were hMPV positive by real time RT-PCR across the different age groups. However, swabs collected from males 35 – 49 years and all patients 50 years and above were negative for hMPV. It is notable that high proportions (7.0%) of swabs were positive for hMPV among the males 18 – 34 years. The hMPV was most common among males 12 – 23 months (8.2%) which suddenly drops in males 24 – 59 months (5.6%) as compared to females in these two age groups where similar proportions were maintained at 7.7% and 7.6% respectively. The lowest proportion was in patients above 50 years old (0.0%).

4.3. Lwak and Kibera incidence rates per 1000 person years

Incidence rates per 1000 person-years of observation (pyo) for both study sites as from 1st October 2006 to 30th September 2009 were analyzed by age group and sex (Table 4.3 and 4.4). Statistically, there is no difference within the males and females from each study site. Overall, the incidence rates per 1000 person-years of observation were significantly higher from the Kibera than from Lwak study site (p-value = 0.0002). In Lwak, females were reported as having higher incidence rates than males as opposed to Kibera, where males had the higher incidence rate compared to females (Table 4.3.)

Table 4.3: Incidence rates of hMPV for all age groups by sex from Lwak and Kibera.

	LWAK			KIBERA		
Gender	***pyo**	**hMPV +**	**hMPV †IR/1000 pyo**	***pyo**	**hMPV +**	**hMPV †IR/1000 pyo**
Male	35,502.19	20	0.56	41,803.66	49	1.17
Female	39,768.37	23	0.58	40,045.30	43	1.07
TOTAL	75,270.56	43	0.57	81,848.96	92	**1.12**

*pyo = person years of observation

†IR = Incidence rate

The incidence rates per 1000 person-years were higher in children < 5 years old than in persons ≥ 5 years old including adults (Table 4.4). In Kibera urban informal settlement, the children less than 5 years old were 2.5 times (RR= 2.46; p-value = 0.0003) more likely to be infected by hMPV than those in Lwak rural community while for older children more than 5 years old and adults, there was no significant differences in the rates of hMPV at both sites (RR=1.26, p-value = 0.3802) (Table 4.4).

Table 4.4: Incidence rates of hMPV in relation to two distinct age groups.

	LWAK			KIBERA		
Age Group	***pyo**	**hMPV +**	**hMPV †IR/1000 pyo**	***pyo**	**hMPV +**	**hMPV †IR/1000 pyo**
< 5 years	11,998.74	19	**1.58**	15,396.43	60	**3.90**
≥ 5 years	63,271.82	24	0.38	66,452.53	32	0.48

*pyo = person years of observation

†IR = Incidence rate

The incidence rates per 1000 person-years from both study sites were highest in males than females less than 5 years old but incidences in Kibera for both sexes < 5 years old was similar (Table 4.5). In the ≥ 5 years old, the females from Lwak had a higher incidence rate per 1000 person-years of observation as compared to Kibera, where the males had the highest incidence rate per 1000 person-years of observation.

Table 4.5: Incidence rates for children and adults by sex for both study sites.

	LWAK			KIBERA		
Gender/ Age Group	***pyo**	**hMPV +**	**hMPV †IR/1000 pyo**	***pyo**	**hMPV +**	**hMPV †IR/000 pyo**
Male < 5	5,986.32	12	**2.00**	7,687.85	30	**3.90**
Female < 5	6,012.42	7	1.16	7,708.58	30	3.89
Male ≥ 5	29,515.87	8	0.27	34,115.81	19	0.56
Female ≥ 5	33,755.95	16	0.47	32,336.72	13	0.40

*pyo = person years of observation

†IR = Incidence rate

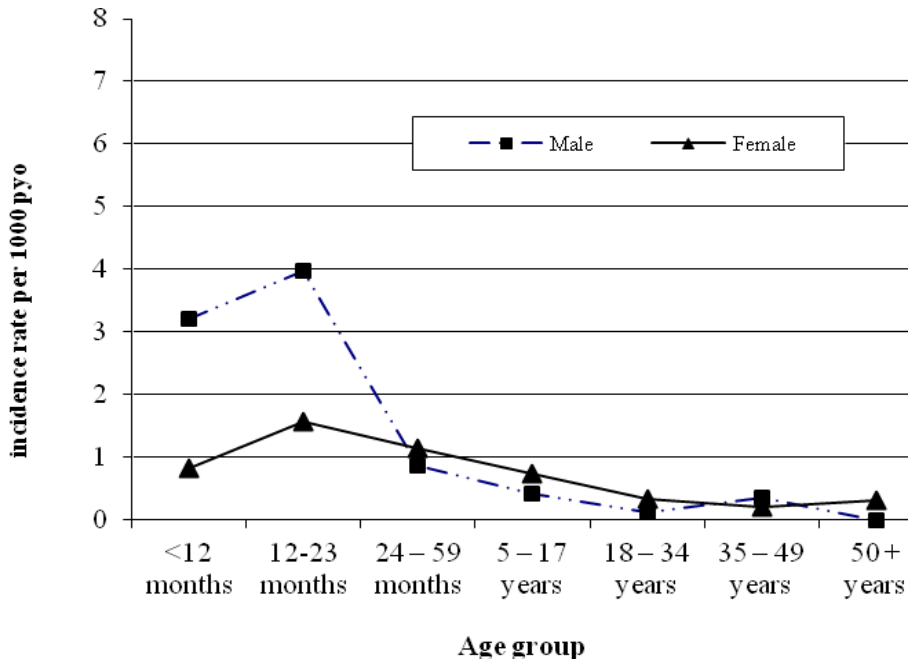


Fig 4.5: Incidence rates of hMPV per 1000 person year for Lwak by age group and sex (n = 43).

The incidence rates expressed per 1000 person years were analyzed across the different age groups by sex for Lwak rural community (Figure 4.5). Incidence rate is high among males up to 59 months than females of the same age group and then becomes lower than females up to 35 years. At 50 years and above, the female incidence rates were higher than for the males of the same age group.

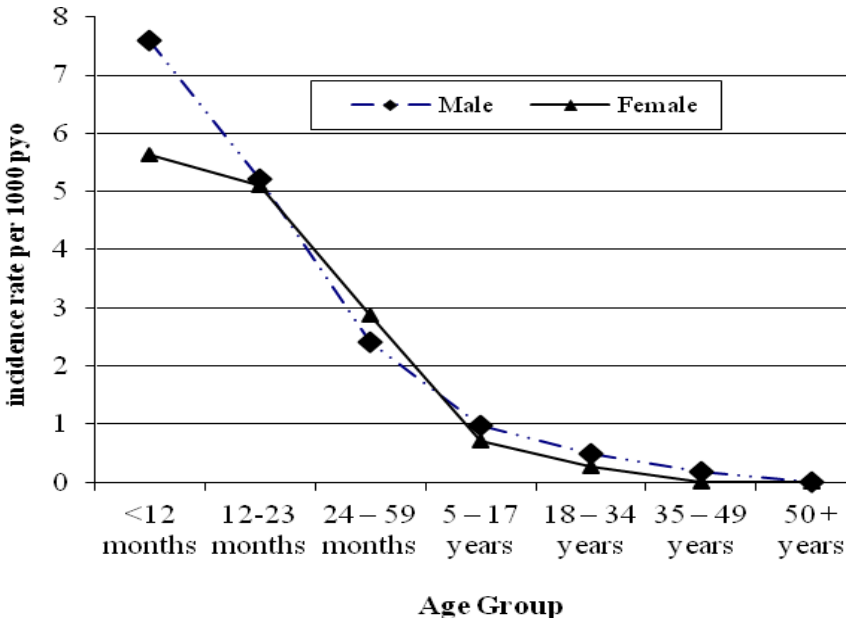


Fig 4.6: Incidence rates of hMPV per 1000 person years for Kibera by age group and sex (n = 92).

The incidence rates for Kibera were analyzed per 1000 person years across the different age groups by sex (Figure 4.6). Incidence rates among males were higher than females up to 23 months. After this, the trends are similar up to 50 years old for both sexes.

4.4. Adjusted incidence rates of hMPV by age group

Incidence rates of hMPV among the study participants were adjusted to cater for patients who were symptomatic but did not attend the specified field clinic and also those who had ARTI but whose samples were not collected (Table 4.6 and 4.7).

Adjusted incidence rates were highest in < 12 months children from Lwak at 29.5/1000 person-years of observation with the lowest among 18 – 34 years old at 0.78/1000 person-years of observation. The overall adjusted incidence rates for Lwak was 4.66/1000 person-years of observation. Adjusted incidence rates were high in children < 5 years old at 27.90/1000 person-years of observation as compared to adults at 1.4/1000 person-years of observation.

Table 4.6: Adjusted incidence rates of hMPV for Lwak, 2006 – 2009

Age Group	Estimated Population under Study	*pyo	Lwak hMPV +	Lwak hMPV †IR/ 1000 pyo	% cases with ALRI sample	Adjusted Rate1∞	% cases visiting field clinic	Final Adjusted †IR[µ]	95% CI
<12m	792	2460.85	5	2.03	13	15.6	44	**29.5**	5.4 – 82.6
12-23m	923	2539.88	7	2.76	13	21.2	47	**27.7**	12.2 – 81.8
24 - 59m	2,503	6998.01	7	1.00	10	10.0	49	**20.4**	12.7 – 85.2
5 – 17yr	10,081	28001.62	16	0.57	28	2.0	47	**5.96**	2.3 – 7.0
18 – 34yr	6,868	17267.56	4	0.23	39	0.6	50	**0.78**	0.0 – 1.0
35 – 49yr	2,581	7645.30	2	0.26	37	0.7	46	**0.80**	-1.7 – 1.1
>50 yr	3,546	10357.34	2	0.19	42	0.5	45	**0.93**	-1.7 – 1.2
TOTAL	27,294	75,261.56	43	0.57	21	2.7	45	**4.66**	3.2 – 5.8
< 5yrs	4,218	11,998.74	19	1.58	12	13.2	43	**27.90**	23.7 – 62.3
Adults	23,076	63,271.82	24	0.38	32	1.19	47	**1.40**	0.3 – 6.6

In Kibera, the highest adjusted incidence rate was also from children < 12 months at 20.0/1000 pyo with the lowest at > 50 years old at 0.0/1000 person-years of observation (Table 4.7). These rates reduced by increase in age. The overall adjusted incidence rates for Lwak was 3.7/1000 person-year of observation. Adjusted incidence rates were high in children < 5 years old at 7.6/1000 person-years of observation.

Table 4.7: Adjusted incidence rates of hMPV for Kibera, 2006 – 2009

Age Group	Estimated Population under Study	*pyo	Kibera hMPV +	Kibera hMPV †IR/ 1000 pyo	% cases with ALRI sample	Adjusted Rate1∞	% cases visiting field clinic	Final Adjusted †IRµ	95% CI
<12m	830	2,426.35	16	6.59	47	14.0	70	**20.0**	12.3 - 32.6
12-23m	1,068	3,879.19	20	5.16	44	11.7	73	**16.0**	10.3 - 24.8
24 - 59m	3,136	9,090.88	24	2.64	37	7.1	78	**9.1**	6.1 - 13.6
5 – 17yr	8,934	23,992.76	20	0.83	76	1.1	77	**1.4**	0.9 - 2.2
18 – 34yr	9,366	30,022.20	11	0.37	60	0.6	70	**0.9**	0.5 - 1.6
35 – 49yr	3,342	9,378.38	1	0.11	53	0.2	68	**0.3**	0.0 - 2.1
>50	797	3,059.13	0	0.00	37	0.0	71	**0.0**	0.0 - 0.0
TOTAL	27,473	81,848.90	92	1.12	41	2.7	73	**3.7**	3.0 - 4.5
< 5yrs	5,034	15,396.42	60	3.90	68	5.7	75	**7.6**	5.9 - 9.8
Adults	22,439	66,452.47	32	0.48	47	1.0	72	**1.4**	1.0 - 2.0

*pyo = person years of observation

†IR = Incidence rate

∞ Adjustment 1 based on percentage of patients who met sample collection criteria who had a sample taken.

µ Adjustment 2 (total adjustment) based on the patients with ALRI during home visit, who visited the field clinic.

4.5. Risk Ratios of hMPV by age groups

Patients who tested positive for hMPV were analyzed for Risk Ratio without reference to the sex (Table 4.8). Age group was used to determine the risk of getting infection by hMPV in relation to a multiplying factor. Age groups 18 – 34 years of the young adults who are presumed to have acquired immunity over time were used as the reference group for analysis of risk to acquire infection.

Children less than 5 years old were significantly (p-value < 0.001) at a higher risk of getting infection by hMPV especially in the urban setting. The children below 5 years in Lwak were 6.8 times more likely to get hMPV as compared to the reference age group and those from Kibera were 10.6 times more likely to be infected by hMPV in relation to the specific reference age group of each location. This study indicated that persons above 50 years were at a low risk of infection by hMPV.

Table 4.8: Lwak and Kibera Risk Ratios for acquiring hMPV

	LWAK			KIBERA		
Age Group	**hMPV IR/ 1000 pyo**	**Risk Ratio**	**p-value (95%CI)**	**hMPV IR/1000 pyo**	**Risk Ratio**	**p-value (95%CI)**
<12 months	2.03	8.77	*<0.001 (2.55 - 30.15)	6.59	**18.00**	*<0.001 (8.50 - 38.10)
12-23 months	2.76	**11.90**	*<0.001 (3.72 - 38.03)	5.16	14.07	*<0.001 (6.85 - 28.91)
24 – 59 months	1.00	4.32	*0.011 (1.35 - 13.80)	2.64	7.21	*<0.001 (3.58 - 14.50)
5 – 17 yrs	0.57	2.47	0.094 (0.87 - 7.03)	0.83	2.28	*0.027 (1.11 - 4.67)
18 – 34 yrs	**0.23**	**REF**	**REF**	**0.37**	**REF**	**REF**
35 – 49 yrs	0.26	1.13	0.888 (0.24 - 5.27)	0.11	0.29	0.211 (0.05 - 1.75)
50 + yrs	0.19	0.83	0.826 (0.18 - 3.89)	0.00	0	0.287 (0.00 - 3.43)

4.6. Clinical symptoms in hMPV patients, Lwak and Kibera

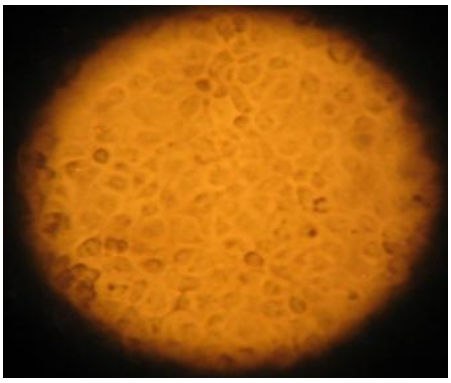
There were varied clinical symptoms for the hMPV positive cases who had symptoms recorded, 93/135 i.e. 69% (Figure 4.7). The most common symptoms in hMPV patients included cough, fever and runny nose, sneezing, Temp $\geq 38^{\circ}C$ and Oxygen Saturation ≤ 90. The remaining symptoms were less common (< 40%) in all patients.

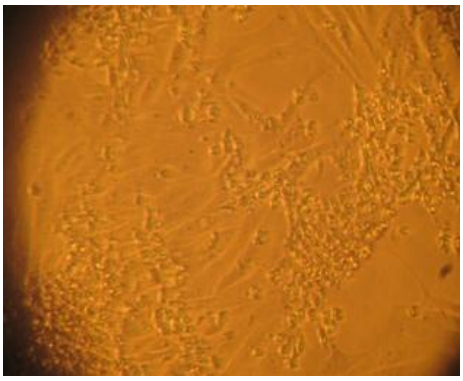
Table 4.9: Common clinical symptoms found in patients with hMPV, 1st October 2006 to 30th September 2009

	Patients < 5 yrs				**Patients ≥ 5 yrs**					**TOTAL (%)**
Symptom	***N**	**<12m**	**12-23m**	**24-59m**	***N**	**5-17y**	**18-34y**	**35-49y**	**50+y**	
Patient with symptoms	**51**	11	19	21	**42**	28	9	3	2	**93 (100)**
Cough	**47**	10	18	19	**37**	23	9	3	2	**84 (90)**
Fever	**46**	9	18	19	**38**	24	9	3	2	**84 (90)**
Runny nose	**41**	8	14	19	**35**	23	7	3	2	**76 (82)**
Sneezing	**29**	6	13	10	**30**	20	6	3	1	**59 (63)**
Temperature ≥38	**26**	8	7	11	**25**	17	4	2	2	**51 (55)**
Oxygen Saturation ≤90	**31**	5	12	14	**20**	14	5	1	0	**51 (55)**
Difficulty in breathing	**18**	5	10	3	**12**	5	5	2	0	**30 (32)**
Vomits	**14**	4	5	5	**10**	5	4	1	0	**24 (26)**
Sore throat	**5**	0	3	2	**14**	11	3	0	0	**19 (20)**
Chest pain	**3**	1	2	0	**15**	6	7	1	1	**18 (19)**
Lethargic	**4**	0	2	2	**2**	2	0	0	0	**6 (6)**
Cannot drink	**0**	0	0	0	**2**	2	0	0	0	**2 (2)**
Stridor	**0**	0	0	0	**0**	0	0	0	0	**0 (0)**

4.7. Subtypes of hMPV in Kenya

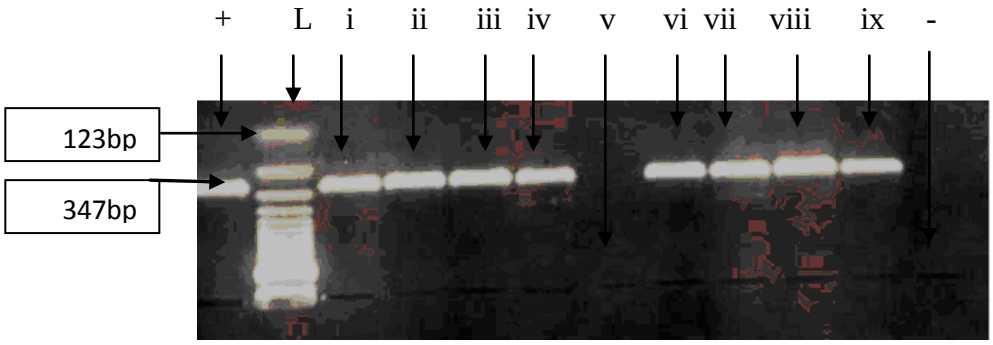
Cytopathic effects (CPE) were clearly visible in 17 out of 30 isolates and this was seen usually after the 16^{th} day (Plate 4.1). CPE presented with rounding of the cells and later degeneration of the cell monolayer (Plate 4.1). The flasks were then stored at -80˚C when about 80% had shown CPE. The remaining 13 were either contaminated by bacteria and/or fungi while others did not show CPE or turn positive on confirmation by rt RT-PCR. Isolation of hMPV was attempted in 30 NP/OP specimens i.e.18 from Lwak Mission Hospital and 12 from Kibera. The virus was successfully recovered from a total of 17 specimens (57%). The primer set used for RT-PCR could detect all hMPV lineages based on the F-gene (Plate 4.2). These were later differentiated into the subtypes by DNA sequencing. Thirteen isolates of hMPV from Lwak and 4 isolates from Kibera were characterized. There was no hMPV isolated that could not be subtyped using RT-PCR and DNA sequencing. CPE was not observed in isolate 2006772679KIB (v) but was subjected to RT-PCR to rule out the presence of hMPV.





(i) Magnification X100 (ii) Magnification X100

Plate 4.1: Picture indicating uninfected (i) and infected (ii) LLC-MK2 cells with hMPV respectively



Key: Lanes + =**Positive control**, L= Molecular Marker, i = 2008762909ABI, ii = 2007752345BON, iii = 2007738983NYA, iv = 2007752154BON, v = 2006772679KIB, vi = 2007752023ABI, vii = External Isolate for PCR confirmation, viii = 2009791118KIB, ix = 2007753069BON and - = **Negative control**

ABI = Abidha, BON = Bondo, NYA = Nyangoma from Lwak study site and KIB = Kibera

Plate 4.2: Picture of agarose gel stained with Ethidium Bromide indicating RT-PCR positives for hMPV

4.8. Nucleotide sequence identity

Generated sequences were analyzed for percent average identity of hMPV sequences to determine the percent isolate identity in comparison with each of the other isolates (Table 4.10). The mean percent nucleotide identity of all cultured isolates was determined with 100% identity indicating the same isolate. Isolate nucleotide identity was > 90% for each of the three determined subgroups i.e. A2, B1 and B2. The average pairwise nucleotide similarity among all the 17 isolates from Kenya was 88.0% (Table 4.11). No subtype A1 was identified with five isolates being A2, two B1 and ten B2 with average nucleotide similarity within the different subtypes being 93.9%, 97.7% and 95.7% respectively. There was no sequence pattern indicating a difference between the 4 isolates from Kibera, which had an average pairwise nucleotide similarity of 85.7 % when compared to the 13 isolates from Lwak that had an average nucleotide pairwise similarity of 89.2 %.

Table 4.10: The nucleotide percentage identity of the isolated hMPV sequences

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		Av % ID
1	2007753043BON	100	92.0	80.4	77.7	99.3	96.7	99	99.3	80.7	78.7	98.7	98.7	80.1	91.7	94.7	92.0	98.7	94.0	81.4	82.1	73.4	92.0	97.0	1	91.3
2	2007752303BON		100	82.4	79.1	92.0	89.4	91.7	92.0	82.7	80.7	91.4	91.4	82.1	89	87.7	98.7	92.0	90.0	82.7	84.1	73.4	97.3	93.0	2	88.3
3	2007753070BON			100	95.7	80.4	78.4	80.1	80.1	99.3	97.7	79.7	80.4	98.3	77.1	76.4	82.7	80.4	78.1	93.4	94.7	71.8	81.7	80.4	3	84.3
4	2007753069BON				100	77.7	75.7	77.7	78.4	96.0	94.4	77.7	78.7	94.7	75.7	76.7	79.0	78.7	75.4	89.7	91.0	68.8	78.4	77.1	4	81.2
5	2008762902ABI					100	96.7	99.0	99.3	80.7	78.7	99.3	98.7	80.1	91.4	94.7	92.0	98.7	94.7	81.4	82.1	73.4	92.0	97.0	5	92.6
6	2009748064LWA						100	96.3	96.7	78.7	76.7	96.3	96.0	78.1	88.7	92.0	89.3	96.0	91.4	79.4	80.1	71.4	89.4	94.4	6	89.7
7	2009743201LWA							100	99.3	80.4	78.4	98.7	98.7	80.1	91	95.0	91.7	98.7	94.0	81.1	81.7	73.8	91.7	96.7	7	91.5
8	2007753048BON								100	80.8	78.7	99.3	99.3	80.1	91.4	95.0	92.0	99.3	94.7	81.4	82.1	73.4	92.0	97.0	8	91.1
9	2007753039BON									100	98.0	80.1	80.7	98.7	77.4	76.7	82.7	80.7	78.4	93.7	95.0	71.8	82.1	80.7	9	83.7
10	2006772773KIB										100	78.1	78.7	97.3	75.4	75.7	80.7	78.7	76.4	92.4	93.7	71.1	80.4	78.7	10	80.1
11	2009791090KIB											100	98.7	79.4	90.7	94.4	91.3	98.7	94.7	80.7	81.4	73.4	91.4	96.3	11	92.6
12	2009791118KIB												100	80.1	91.0	94.4	91.3	99.3	94.0	81.4	82.1	72.8	91.4	96.3	12	91.7
13	2009790864KIB													100	76.7	76.1	82.0	80.1	77.7	93.0	95.0	71.8	81.4	80.1	13	78.5
14	2008762909ABI														100	89.0	88.3	91.0	93.4	78.4	79.4	70.8	89.0	93.7	14	90.4
15	2007752345BON															100	88.0	94.4	89.7	77.4	78.1	71.4	88.4	92.4	15	90.7
16	2007738983NYA																100	92.0	90.0	82.0	83.3	74.0	98.0	93.0	16	91.0
17	2007752154BON																	100	94.0	81.4	82.1	73.1	91.4	96.3	17	94.0
18	IRH1207																		100	79.7	80.4	72.4	90.0	95.7	18	
19	A1																			100	95.0	74.1	82.1	81.7	19	
20	A2																				100	73.1	83.4	82.7	20	
21	AMPV_C																					100	72.8	73.1	21	
22	B1																						100	93.0	22	
23	B2																							100	23	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		
	Av % ID		92	81.4	84.17	87.35	87.38	90.63	92.16	84.91	84.67	89.93	90.91	85.76	85.17	87.04	88.11	91.09	88.27							88.0

Key: Green = A2 subtypes **Black = B1 subtypes** **Red = B2 subtypes**

Table 4.11: Mean percent identities in relation to the hMPV subtypes

Subtypes	No. of sequences	Minimum % nt identity	Mean % nt identity
A2	5	91	93.9
B1	2	97.3	97.7
B2	10	92.4	95.7
All	17	91	**88.0**

Key:

% = percent

No. = number

nt = nucleotide

Phylogenetic analysis showed that 5 of the 17 (29 %) Kenya viruses belonged to the A subtype, whereas the remaining 12 (71%) viruses belonged to the B subtype. The B group were further subdivided into B1 subtypes (2 strains) and B2 strains subtype (10 strains). All the A subtypes were A2 subtype and no A1 subtypes were identified from the isolates. The minimum nucleotide percent identity was 91% in the A2 subtype and approximately 98% in the B1 subtype (Table 4.11).

4.9.Phylogenetic analysis of the hMPV isolates

FigTree v 1.2.2 software (Rambaut, 2006) was used to view the phylogenetic tree of the 17 isolates to review the genetic relatedness (Figure 4.7). Isolates that are of similar identity (i.e. closer to 100% nucleotide identity) are known to cluster together. It was observed that all the isolates clustered with published reference strains, but the isolates deviated from the reference strains. The two distinct hMPV groups were clearly divergent with 100% bootstrap values.

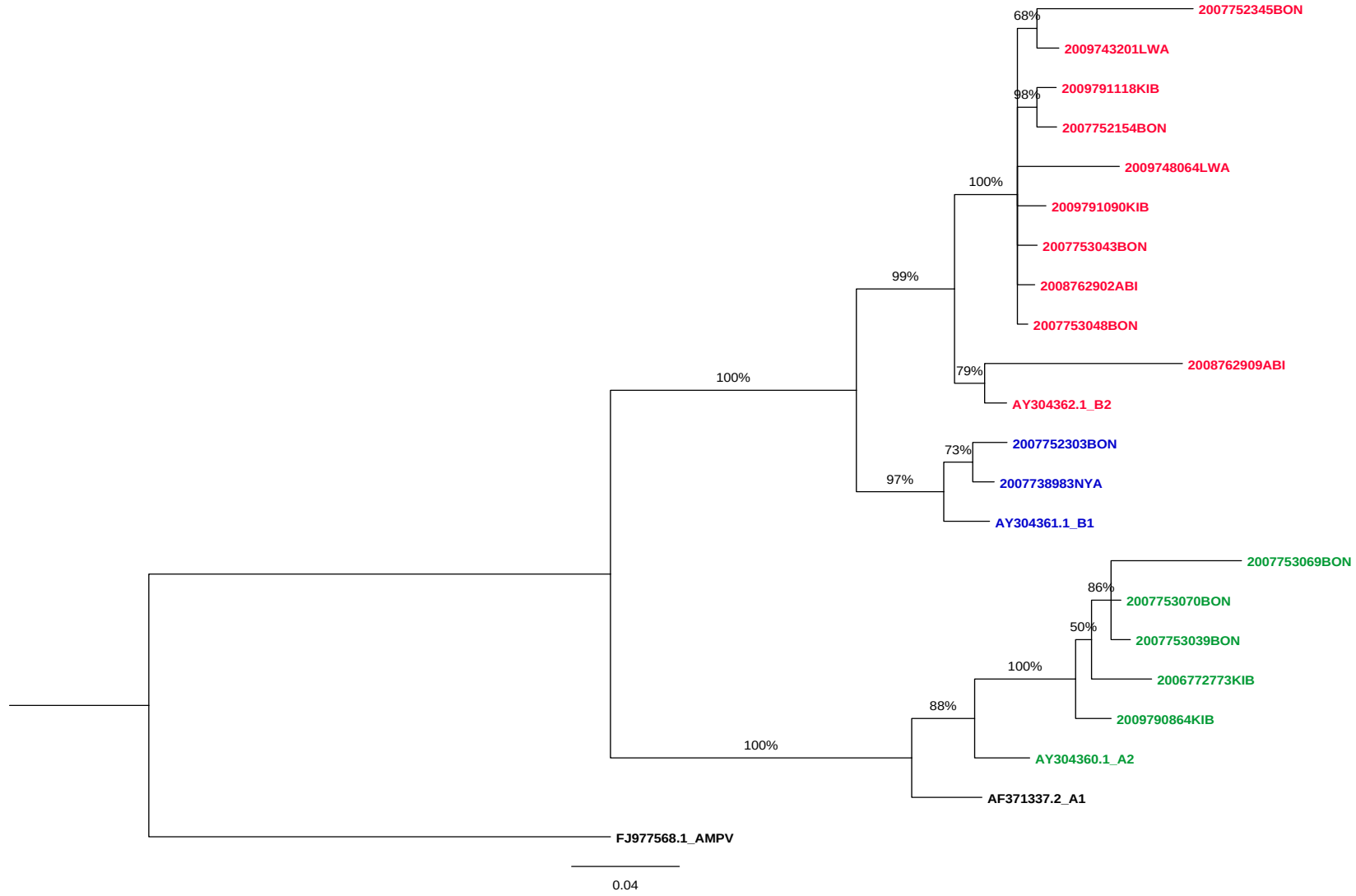


Fig 4.7: Phylogenetic tree showing subtypes of hMPV isolates from the selected populations

CHAPTER FIVE

5.0. DISCUSSION

It is possible that this is the first comprehensive study attempting to determine the incidence rates of hMPV and to characterize the circulating subtypes of the virus in Kenya. *Metapneumovirus* has been described worldwide associated with a considerable respiratory morbidity especially in children < 5 years old (Boivin *et al.,* 2004). In the original study by van den Hoogen *et al.,* 2001, more than 96% of infected children were < 5 years old. The hMPV incidence rate from both Lwak and Kibera among children < 5 years old was 2.88/1000 person-years of observation, which decreases to 0.43 per 1000 person-years of observation for children ≥ 5 years old and adults. However, in the < 5 years old age group, the study revealed that the males were more affected. The incidence rate trends were similar among the urban Kibera and rural Lwak population, but Kibera had higher incidence rates than Lwak, probably due to overcrowding and poor living conditions. The trend indicating reducing incidence rates with an increasing age for both males and females in this study is similar to findings in a previous study in Finland (Heikkinen *et al.,* 2008). Despite the fact that the study population had a high percentage of individuals at the 5 – 34 years age bracket (62 – 67%), the highest incidence rates of hMPV was seen in children < 2 years which is in agreement with previous studies (Williams *et al.,* 2004; Garcia *et al.,* 2006; Heikkinen *et al.,* 2008).

Adjusted incidence rates were higher in children < 5 years from Lwak rural community (27.9/1000 pyo) than those from Kibera urban informal settlement (7.6/1000 pyo). However, persons ≥ 5 years old from both study sites had the same adjusted incidence rate of 1.4/1000 pyo. This age category may not be affected by variation in population density. The swabs collected indicate the presence of hMPV in all age categories. However, swabs collected from children less than 5 years old from both study sites indicate a higher proportion of hMPV positives than in children above 5 years old, adults and the elderly. This could be an indication that the hMPV is a common etiology among children < 5 years old. Therefore, the possibility of isolating and characterizing hMPV is greater in samples obtained from children < 5 years old than from persons ≥ 5 years old in a general population. This is comparable to other studies that have primarily focused on determining existence of the hMPV in children alone (Mullins *et al.,* 2004; Madhi *et al.,* 2007). The study showed that hMPV from both study sites is high in males 12 – 23 months old and then a sudden drop in the 24 – 59 months age category. This is different for females of the same age category, where the positive rates rise from the 12 – 23 months to 24 – 59 months in Lwak while in Kibera the rate of positivity is maintained. The incidence rates of hMPV are higher in male children < 2 years which is comparable to the hMPV positivity proportions in this study. In contrast, the study also revealed that a high proportion of swabs collected from males 18 – 34 years old in Kibera had a high proportion of the virus indicating high transmission in this age group possibly due to high population density.

Notably, the rates of hMPV positives are also high between ages 5 – 17 year old in this study. This age bracket includes the majority of school-going children and may be associated with increased transmission rates.

In this study, the hMPV infection was attributed to 2.7% of all ARTI cases from Lwak and 5.1% ARTI cases from Kibera from all age groups. Previously, other studies found the hMPV infection as 3.5% in children < 13 years old and up to 7% in outpatient adults (Osterhaus *et al.,* 2003; Heikkinen *et al.,* 2008). With advances in molecular biology technology worldwide and in developing countries, more etiologies including respiratory pathogens can now be identified and characterized.

Despite the hMPV incidence rates decreasing with an increase in age, the elderly (> 50 years) become vulnerable to the virus infection (Gaunt *et al.,* 2009) similar to this study where the elderly are also affected especially from the Lwak rural community. This could be associated with decreasing immunity in the elderly. Studies indicate that by the time a child is ≥ 5 years, >96 % will have been infected by the hMPV virus (van den Hoogen *et al.,* 2001; Kahn, 2006). Risk Ratio for acquiring hMPV for the study sites were up to 10.6 times higher (p-value < 0.0001) in children < 5 years old than people aged 18 – 34 years in Kibera which is in an urban setting where housing is crowded as opposed to the rural set up of Lwak where the children < 5 years were 6.8 times higher (p-value < 0.0001).

The clinical symptoms mostly associated with hMPV infection in the study were cough, fever, runny nose, sneezing, temperature ≥ 38˚C and oxygen saturation ≤ 90.

In contrast, wheezing and stridor were not observed in hMPV patients from these two study sites. However, other reports have shown that hMPV infection is associated with stridor, cough, rhinitis, fever, laryngitis, nasal congestion, shortness of breath and oxygen saturation ≤ 90 (Heikkinen *et al.,* 2008; Mullins *et al.,* 2004).

Phylogenetic analysis of the hMPV isolates and nucleotide sequence identity revealed the existence of the two major genetic lineages that were strongly supported by high posterior probabilities or bootstrap values (100%) as shown in this study compared to others (Ludewick *et al.,* 2005; Huck *et al.,* 2006;). The two major genetic lineages of the hMPV isolates from these study populations are comparable to subtypes circulating around the world. Majority (59%) of the Kenyan isolates grouped with subtype B2. Subtype A1 was not seen from the isolates probably due to predominant subtypes circulating during alternate seasons. The virus has previously been shown to have one predominant genotype occurring in one season and later having a different genotype in subsequent seasons. An alternation at the subtype level is also evident in subsequent seasons possibly due to genetic shift (Agapov *et al.,* 2006). The 88% identity among all the isolates was similar to previously described identities of 89% over 20 year period (Yang *et al*., 2009) and also 80% overall nucleotide identity (Biacchesi *et al.,* 2003).

Isolation of hMPV in LLC-MK2 cells in this study required up to 21 days for visibility of CPE.

Similar findings have been reported in other studies indicating that hMPV has poor *in vitro* replication kinetics in different cell lines while isolation in cells may also cause amino acid changes due to cell culture adaptation (Crowe *et al.*, 2004; Deffrasnes *et al.*, 2005; Ludewick *et al.*, 2005). However, isolation of a virus is important for further molecular characterization of the complete genome, analyzing pathogenesis in animal models and vaccine development with genetic analysis guiding in studies on transmission rates. The viruses isolated from the two populations were similar to viruses identified elsewhere in the world since all isolates obtained could be subtyped suing DNA sequencing and clustered well with published reference strains.

The variability of hMPV genotypes poses a challenge in vaccine development which relies on molecular and epidemiologic data. One group or genotype and/or subtype can eventually dominate with varied seasons with a possibility of hMPV being the leading cause of respiratory infections within a selected population (Ludewick *et al.*, 2005, Mackay *et al.*, 2006).

CHAPTER SIX

6.0. CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

- These results show that hMPV is present in the selected Kenyan populations and is a cause for acute respiratory illnesses among all age groups but significant in children less than 5 years old (p-value < 0.001) resulting in hospital visits. However, incidence rates of hMPV decline with an increase in age. However, in this age group, the children < 2 years old are at a greater risk of acquiring the infection. Previously, minimal studies have been carried out in developing countries where diagnosis has been a challenge both in technical expertise and expense.

- Majority of the clinical symptoms associated with acute respiratory tract infections were observed from the study participants who had their symptoms recorded. The clinical symptoms mostly associated with hMPV infection in the study were cough, fever, runny nose, sneezing, temperature $\geq 38^{\circ}C$ and oxygen saturation ≤ 90. However, stridor was not observed in the hMPV cases as compared to previous studies.

- Group A and B genetic lineages of hMPV are in circulation within the study populations. Nucleotide identity of the isolates is similar to previously designed studies.

6.2. RECOMMENDATIONS

- Monitoring of seasonal trends and continued surveillance to assist in analysis of risk groups and establish measures that can be put in place to reduce the morbidity and mortality associated with hMPV as a contributor to the respiratory disease burden. Simpler and affordable diagnostic methods e.g. Oxoid Imagen hMPV ELISA should be introduced in developing countries to further enhance surveillance and assist in instituting control measures.

- Further analysis of clinical symptoms associated with acute respiratory tract infections to increase sensitivity of getting cases. Seroprevalence studies should be implemented to highlight the impact hMPV infection among the Kenyan population. Economic analysis studies should be carried out to highlight the importance of respiratory diseases in relation to clinic visits and costs of medication on the affected population.

- Further subtyping and DNA sequencing of the whole genome of hMPV within the Kenyan population is highly advisable to fully characterize the circulating strains and fill the gap in knowledge in relation to vaccine development for the developing countries. Referring to the long periods it takes to isolate the hMPV virus in a specialized continuous cell line, genetic analysis can be carried out directly from the specimens to reduce the impact that may result from mutation.

REFERENCES

Agapov, E., Sumino, K. C., Gaudreault-Keener, M., Storch, G. A. and Holtzman, M. J. (2006), Genetic variability of human metapneumovirus infection: evidence of a shift in viral genotype without a change in illness, *J. Infect. Dis.* 193:396–403.

Al-Sonboli, N., Hart, C. A., Al-Aghbari, N., Al-Ansi, A., Ashoor, O. and Cuevas, L. E. (2006), Human metapneumovirus and Respiratory Syncytial Virus Disease in Children, Yemen, *Emerg Infect Dis.*12:9; 1437 – 1439.

Alvarez, R., Lwamba, H. M., Kapczynski, D. R., Njenga, M. K. and Seal, B. S. (2003), Nucleotide and predicted amino acid sequence-based analysis of avian metapneumovirus type C cell attachment glycoprotein gene; phylogenetic analysis and molecular epidemiology of U. S. pneumovirus, *J. Clin. Microbiol.* 41;4; 1730 – 1735.

Barenfanger, J., Mueller, T., O'Brien, J., Drake, C. and Lawhorn, J. (2008), Prevalence of Human metapneumovirus in Central Illinois in patients thought to have respiratory viral infections, *J. Clin. Microbiol.* 46; 4; 1489–1490.

Bastien, N., Ward, D., Van Caeseele, P., Brandt, K., Lee, S. H., McNabb, G., Klisko, B., Chan, E. and Li, Y. (2003), Human metapneumovirus infection in the Canadian population, *J. Clin. Microbiol.* 41; 10; 4642–4646.

Bastien, N., Liu, L., Ward, D., Taylor, T. and Li, Y. (2004), Genetic variability of the G glycoprotein gene of human metapneumovirus, *J. Clin. Microbiol.* 42; 8:3532–3537.

Biacchesi, S., Skiadopoulos, M. H., Boivin, G., Hanson, C. T., Murphy, B. R., Collins, P. L. and Buchholz, U. J. (2003), Genetic diversity between human metapneumovirus subgroups, *Virology* 315**:**1–9.

Black, R. E., Morris S.S. and Bryce J. (2003), Where and why are 10 million children dying every year? *Lancet;* 361:2226-2234.

Boivin, G., Mackay, I., Sloots, T. P., Madhi, S., Freymuth, F., Wolf, D., Shemer-Avni, Y., Ludewick, H., Gray, G. C. and LeBlanc, E. (2004), Global genetic diversity of human metapneumovirus fusion gene, *Emerg. Infect. Dis.*10:6:1154–1157.

Brooks, W. A., Santosham, M., Neheed, A., Gowswami, D., Wahed, M. A., Diener-West, M., Faruque, A. S. G. and Black, R. E. (2005), Effects of weekely zinc supplements on incidence of pneumonia and diarrhoea in children younger than 2 years in an urban, low-income population in Bangladesh: randomized controlled trial, *Lancet;* 366: 999 – 1004.

Broor, S. and Bharaj, P. (2007), Avian and Human metapneumovirus, *Ann. N. Y. Acad. Sci.* 1102: 66 – 85.

Buys, S. B. and Du Preez, J. H. (1980), A preliminary report on the isolation of a virus causing sinusitis in turkeys in South Africa and attempts to attenuate the virus, *Turkeys* 28: 36 – 46.

Chano, F., Rousseau, C., Laferriere, C., Couillard, M. and Charest, H. (2005), Epidemiological survey of human metapneumovirus infection in a large pediatric tertiary care center, *J. Clin. Microbiol.* 43:11:5520–5525.

Chung, J. Y., Tae, H. H., Byung, E. K., Kim, C. K., Sang, W.K. and Hwang, E. (2006), Human Metapneumovirus Infection in Hospitalized Children with Acute Respiratory Disease in Korea, *J Korean Med Sci* 21: 838-42.

Crowe, E. C. (2004), Human Metapneumovirus as a major cause of human respiratory tract disease, *The Ped. Inf. Dis J.* 23; 11; S215-S221.

Dare, R., Sanghavi, S., Bullotta, A., Keightley, M. C., St. George, K., Wadowsky, R. M., Paterson, D. L., McCurry, K. R., Reinhart, T. A. ***et al*** (2007), Diagnosis of Human Metapneumovirus Infection in Immunosuppressed Lung Transplant Recipients and Children Evaluated for Pertussis, *J. Clinical Micro,* 45:2:548–552.

Deffrasnes, C., Coˆte, S., and Boivin, G. (2005), Analysis of Replication Kinetics of the Human Metapneumovirus in Different Cell Lines by Real-Time PCR, *J. Clinical Micro*, 43:1: 488–490.

Edgar, R. C., (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucl. Acids Research* 32; 5; 1792 – 97.

Esper, F., Martinello, R. A., Boucher, D., Weibel, C., Ferguson, D., Landry, M. L. and Kahn, J. S. (2004), A 1-year experience with human metapneumovirus in children aged 5 years, *J. Infect. Dis.* 189:8:1388–1396.

Falsey, A. R., Erdman, D., Anderson, L. J. and Walsh, E. E. (2003), Human Metapneumovirus infections in young and elderly adults, *J Infect Dis.* 187:785–90.

Feikin, D. R., Olack, B., Bigogo, G. M., Audi, A., Cosmas, L., Aura, B., Burke, H., Njenga, M. K.,Williamson, J and Breiman, R. F. (2011), The burden of common infectious disease syndromes at the clinic and household level from population-based surveillance in Rural and Urban Kenya, *Plos.* 6; 1; 1 – 10.

Garcia-Garcia, M. L., Calvo, C. and Martin, F. (2006), Human metapneumovirus infections in hospitalized infants in Spain, *Arch Dis Child.* 91; 290 – 295.

Gaunt, E., McWilliam-Leitich, E. C., Templeton, K. and Simmonds, P. (2009), Incidence, molecular epidemiology and clinical presentations of human metapneumovirus; assessment of its importance as diagnostic screening target, *J. Clinic. Virol* 46:4:318 – 324.

Gerna, G., Campanini, G., Rovida, F., Sarasini, A., Lilleri, D., Paolucci, S., Marchi, A., Baldanti, F. and Revello, M. G. (2005), Changing circulation rate of human metapneumovirus strains and types among hospitalized pediatric patients during three consecutive winter-spring seasons, *Arch. Virol.* 150:11**:**2365–2375.

Girard, M. P., Cherian, T., Pervikov, Y. and Kieny, M. P. (2005), A review of vaccine research and development: Human acute respiratory infections, *Vaccine* 23: 5708 – 5724.

Hall, T. (1999), Bioedit software: http://www.mbio.ncsu.edu/BioEdit/bioedit.html

Heikkinen, T., Osterback, R., Peltola, V., Jartti, T. and Vainionpaa, R. (2008), Human Metapneumovirus Infection is children, *Emerg Infect Dis.* 14(1): 101 – 106.

Hopkins, M. J., Redmond, C., Shaw, J. M., Hart, I. J., Hart, C. A., Smyth, R. L. and Semple, M. G. (2008), Detection and characterization of human metapneumovirus from children with acute respiratory symptoms in north-west England, UK, *J. Clinic. Virol.* 42: 273 – 279.

Huck, B., Scarf, G., Haefelin, D. N., Puppe, W., Weigl, J. and Falcone, V. (2006), Novel Human metapneumovirus sublineage, *Emerg Infect Dis.* 12: 1:147 – 150

Jones, R. C. (1996), Avian pneumovirus infections: questions still unanswered, *Avian Pathol.* 25: 639 – 648.

Kahn, J. (2006), Epidemiology of Human Metapneumovirus, *Am Society of Micro* 19:3: 546 – 557.

Kaida, A., Iritani, N., Kubo, H., Shiomi, M., Kohdera, U. and Murakami, T. (2006), Seasonal distribution and phylogenetic analysis of human metapneumovirus among children in Osaka City, Japan, *J. Clinic. Virol* 35: 394 – 399.

Lin, P. Y., Lin, T. Y., Hunag, Y. C., Tsao, K. C. and Huang, Y. L. (2005), Human Metapneumovirus and community-acquired pneumonia in children, *Chang Gung Med J Vol.* 28:10; 683 – 688.

Ludewick, H. P., Abed, Y., van Niekerk, N., Boivin, G., Klugman, K. P. and Madhi, S. A. (2005), Human Metapneumovirus Genetic Variability, South Africa, *Emerg Infect Dis.*11: 7; 1074 – 1078.

Lwamba, H. C. M., Alvarez, R., Wise, M., Yu, Q., Halvorson, D., Njenga M. K. and Seal B. S. (2005), Comparison of the full-length genome sequence of Avian metapneumoviurus subtype C with other paramyxoviruses, *Virus Res.* 107:1:83-92.

Mackay, I. M., Bialasiewicz, S., Jacob, K. C., McQueen, E., Arden, K. E., Nissen, M. D. and Sloots, T. P. (2006), Genetic Diversity of Human Metapneumovirus over 4 Consecutive Years in Australia, *J. Inf Dis.* 193; 1630-1633.

Madhi, S. A., **Ludewick, H., Kuwanda, L., van Niekerk, N., Cutland, C., Klugman, K. P.** (2007), Seasonality, Incidence and repeat Human metapneumovirus Lower Respiratory Tract Infections in an area with high prevalence of Human Immunodeficiency Virus Type-1 infection, *Pediatr Infect Dis J.* 26: 8; 693-699.

McAdam, A. J., Hasenbein, M. E., Feldman, H. A., Cole, S. E., Offermann, J. T., Riley, A. M. and Lieu, T. A. (2004), Human metapneumovirus in children tested at a tertiary-care hospital *J. Infect. Dis.* 190:20–26.

McConnochie, K. M., Hal, C. B. and Barker, W. H. (1988), Lower respiratory tract illness in the first two years of life; epidemiologic patterns and costs in a suburban pediatric practice, *Am J Public Health.* 78: 34-39.

Mirza, N.M., Macharia, W.M., Wafula, E.M., Agwanda,. R. and Onyango, F.E. (1990) Mortality patterns in a rural Kenyan community. *East Afr Med J;* 67:11:823-829.

Mullins, J. A., Erdman, D., Weinberg, G. A., Edwards, K., Hall, C. B., Walker, F. J., Iwane, M. and Anderson, L. J. (2004) Human metapneumovirus infection among children hospitalized with acute respiratory illness, *Emerg Infect Dis.*10: 4; 700 – 705.

Murray, C., Lopez, A., Mathers, C. and Stein, C. (2001), The Global burden of Disease 2000 project: aims, methods and data sources. In Global programme on evidence for health policy, Discussion paper No. 36, 1-57, *WHO, Geneva Switzerland.*

Njenga, M. K., Lwamba, H. M. and Seal, B. S. (2003), Metapneumoviruses in birds and humans, *Virus Res.* 91:163-169.

Noor, A.M., Zurovac, D., Hay, S.I., Ochola, S.A. and Snow, R.W. (2003) Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. *Trop Med International Health:* 8; 10: 917-26.

Osterhaus, A. and Fouchier, R. (2003), Human metapneumovirus in the community, *Lancet* 361:9361:890-891.

Pelletier, G., **Dery P., Abed Y. and Boivin G.** (2002), Respiratory Tract reinfections by the new Human metapneumovirus in an Immuno-compromised child, *Emerg Infect Dis.* 8:9; 976 – 978.

Peret, T. C., Boivin, G., Li, Y., Couillard, M., Humphrey, C., Osterhaus, A., Erdman, D. and Anderson, L. (2002), Characterization of Human Metapneumovirus isolated from patients in North America, *J. Infect Dis.* 185; 1660 – 1663.

Principi, N., Bosis, S. and Esposito, S. (2006), Human Metapneumovirus in paediatric patients, *Clin Microbiol Infect*. 12(4):301-308.

Rambaut, A. (2006), Fig Tree software, http://tree.bio.ed.ac.uk/

Ronquist, F. and Huelsenbeck, J. P. (2003), MRBAYES 3: Bayesian phylogenetic inference under mixed models, *Bioinformatics* 19:12:1572-1574.

Sanger F., Nicklen, S., Coulson, A.R., (1977) DNA sequencing with chain-terminating inhibitors**,** *Proceedings of the National Academy of Sciences of U S A.* 74(12): 5463–5467.

Schildgen O., Glatzel T., Gaikowski T., Sceibner B., Matz B., Bindl L., et al (2005). Human metapneumovirus RNA in encephalitis patient. *Emerg Infect Dis.* 11; 3:467-470.

Schowalter, R. M., Smith, S. E. and Dutch, R. E. (2006), Characterization of human metapneumovirus F-protein promoted membrane fusion: critical roles for proteolytic processing and low pH. *J. Virol.* 80(22); 10931 – 10941.

Scott, J. A. G., Brooks, W. A., Peiris, J. S. M., Holtzman, D. and Mulholland, E. K. (2008), Pneumonia research to reduce childhood mortality in the developing world, *J Clin Invest* 118:4:1291-1300.

Skiadapolous, M. H., Biacchesi, S., Buchholz, U. J., Riggs, J. M., Surman, S. R., Amaro-Carambot, E., McAuliffe, J. M., Elkins, W. R., St. Claire, M. et al. (2004), The two major human metapneumovirus genetic lineages are highly related antigenically, and the fusion (F) protein is a major contributor to this antigenic relatedness, *J. Virol* 78:13; 6927 – 6937.

Tang, R. S., Schickli, J. H., MacPhail, M., Fernandes, F., Bicha, L., Spaete, J., Fouchier, R. A. M., Osterhaus, A. D., Spaete, R et al., (2003), Effects of Human metapneumovirus and respiratory syncytial virus antigen insertion in two 3' proximal genome positions of Bovine/Human parainfluenza virus type 3 on virus replication and immunogenicity, *J. Virol*. 77:20; 10819 - 10828

Tavare, S. (1986), Some probabilistic and statistical problems on the analysis of DNA sequences *Amer Math Soc: Lect. Math. Life Sci.* 17:57-86.

Tornheim, J. A., Manya, A. S., Oyando, N., Kabaka, S., Breiman, R. F. and Feikin, D. R. (2007), The epidemiology of hospitalized pneumonia in rural Kenya: the potential of surveillance data in setting public health priorities, *Int. J. Inf Dis.* 11: 536 – 543.

Van den Hoogen, B. G., Bestebroer, T. M., Osterhaus, A. D. and Fouchier, R. A. (2002), Analysis of the genomic sequence of a human metapneumovirus, *Virol.* 295:1:119–132.

Van den Hoogen, B. G., De Jong, J. C., Groen, J., Kuiken, T., De Groot, R., Fouchier, R. A. and Osterhaus, A. D. (2001), A newly discovered human pneumovirus isolated from young children with respiratory tract disease, *Nat. Med.* 7**:**6:719–724.

Viazov, S., Ratjen F., Scheidhauer, R., Fiedler, M. and Roggendorf, M. (2003), High prevalence of Human Metapneumovirus Infection in young children and Genetic Heterogeneity of the viral isolates, *J. Clin. Microbiol.* 41; 7: 3043 – 3045.

Walsh, E. E., Peterson, D. R., and Falsey, A. R. (2008), Human metapneumovirus infections in adults, *Arch Inter Med.* 168 (22): 2489 – 2496.

Williams, J. V., Harris, P. A., Tollefson, S. J., Lisa, B. A., Halburnt-Rush L., Pingsterhaus, J. M., Edwards, K. M., Wright, P. F. and Crowe, J. E. (2004), Human metapneumovirus and lower respiratory tract disease in otherwise healthy infants and children, *N Engl J Med;* 350;5: 443-450.

Williams, J. V., Tollefson, S. J., Nair, S. and Chonmaitree, T. (2006), Association of Human metapneumovirus with acute otitis media, *Int J Pediatr Otorhinolaryngol*: 70(7): 1189–1193

Yang, C., Wang, C. K., Tollefson, S. J., Piyaratna, R., Lintao, L. D., Chu, M., Liem, A., Mark, M., Spaete, R. R. et al. (2009), Genetic diversity and evolution of human metapneumovirus fusion protein over twenty years, *J Virol.* 6:138.

APPENDICES

APPENDIX A: Ethical considerations

This study is part of a larger and on-going KEMRI/CDC study with study number **SSC - 932**, "Active population-based study of major infectious disease syndromes in Western Kenya and Nairobi" which has ethical clearance.

Parents consented for their children. The consent form contents were read to the patient in the language they understood by a translator during the interview. A written consent was obtained from participants who were able to write. Witnessed consent was obtained from participants who were illiterate.

All personnel involved in this study were mandated to adhere to an unwavering code of conduct regarding the confidentiality of patients and patients' information.

APPENDIX B: Morbidity surveillance questionnaire for Kibera and Lwak population-based surveillance study sites

Respiratory morbidity surveillance study questionnaire

Questionnaire number: _____

Date of interview: _____ /_____ / _____Interviewer: _____

Part 1: Demographics

a) Surveillance study patient ID: _____

b) First Name: _____

b) Middle Name: _____

c) Last name: _____

d) Date of Birth: _____ /_____ / _____

e) Age: _____years _____ months

f) Gender: □ Male □ Female

Part 2: Vital Signs

2.0

a) Pulse: _____ **b)** Respiratory: _____

Beats per minute *Breaths per minute*

c) Temperature:____˚C **d)** Oxygen Saturation:_____%

e) Weight_____ (kg) **f)** Blood pressure: ___ /____

2.1

a) Is the patient answering by self? □Yes □ No If "Yes" **go to 2.2**

i. What is the relationship of the person answering?

□Mother □Father

□Babysitter □Maid

□Other relative □Other not related

□Unknown

ii. If child < 5 years old, why can't the patient answer questions by him/herself?

□Patient too sick to answer questions □Patient is a child

□Other, specify_____

2.2. What is the reason for visiting the doctor today?

a) Fever □ how many days? _____

b) Cough □ how many days? _____

c) Shortness of breath □

d) Inability to drink □

e) Vomit □

f) Convulsions □

g) Chest In-drawing □

h) Stridor □

i) Oxygen saturation ≤ 90 □

j) Temperature ≥ 38˚C □

k) Sore throat □

l) Other, specify_____

2.3. Does the patient look severely ill? Yes□ No □

Part 3: Other symptoms

i.

a) Do you have ear problem?

b) Do you have a runny nose?

c) Have you been unable to drink die to this illness?

e) Have you vomited or do you have nausea?

m) In the last month have you travelled outside Nairobi?

Yes□ No□ If "Yes" **Go to Part 3, ii.**

ii. Where did you travel to? _____

Part 4: Sample collection and Diagnostic Test

Tests to be done: □ Yes □No

a) Nasal / Oral Pharyngeal swab? □ Yes □No

Date collected: ____/_____/_____ (dd/mm/yy)

b) Other results: _____

Part 6: Disposition

a) What is the disposition of the child?

□ Home □ Refer to hospital

(Specify)_____

APPENDIX C: Composition of isolation media

1. For Growth Medium:

To a 500 ml Dulbecco's Medium add;

- 5.5 ml of L-Glutamine
- 5.5 ml of Antibiotics (Penicillin + Streptomycin + Amphotericin B)
- 55 ml of Fetal Bovine Serum.
- 0.44 ml of Gentamycin

2. For Maintenance medium:

To a 500 ml of Dulbecco's medium, add;

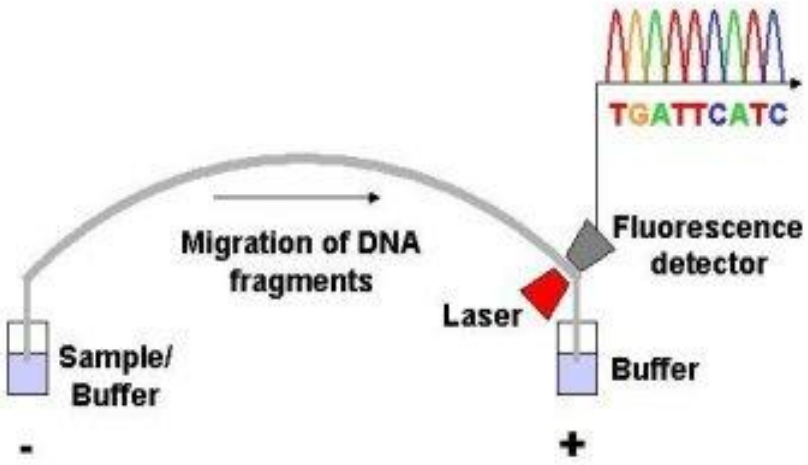
- 5.5 ml of L-Glutamine
- 5.5 ml of Antibiotics (Penicillin + Streptomycin + Amphotericin B)
- 11 ml of Fetal Bovine Serum.
- 0.44 ml of Gentamycin

3. For Viral Growth Medium/Maintenance medium for infected cells:

To a 500ml of Dulbecco's medium, add;

- 5.5 ml of L-Glutamine
- 5.5 ml of Antibiotics(Penicillin +Streptomycin + Amphotericin B)
- 12.5 ml of Bovine Serum Albumin
- 0.44 ml of Gentamycin

APPENDIX D: Operation of the 3130xl Genetic Analyzer.

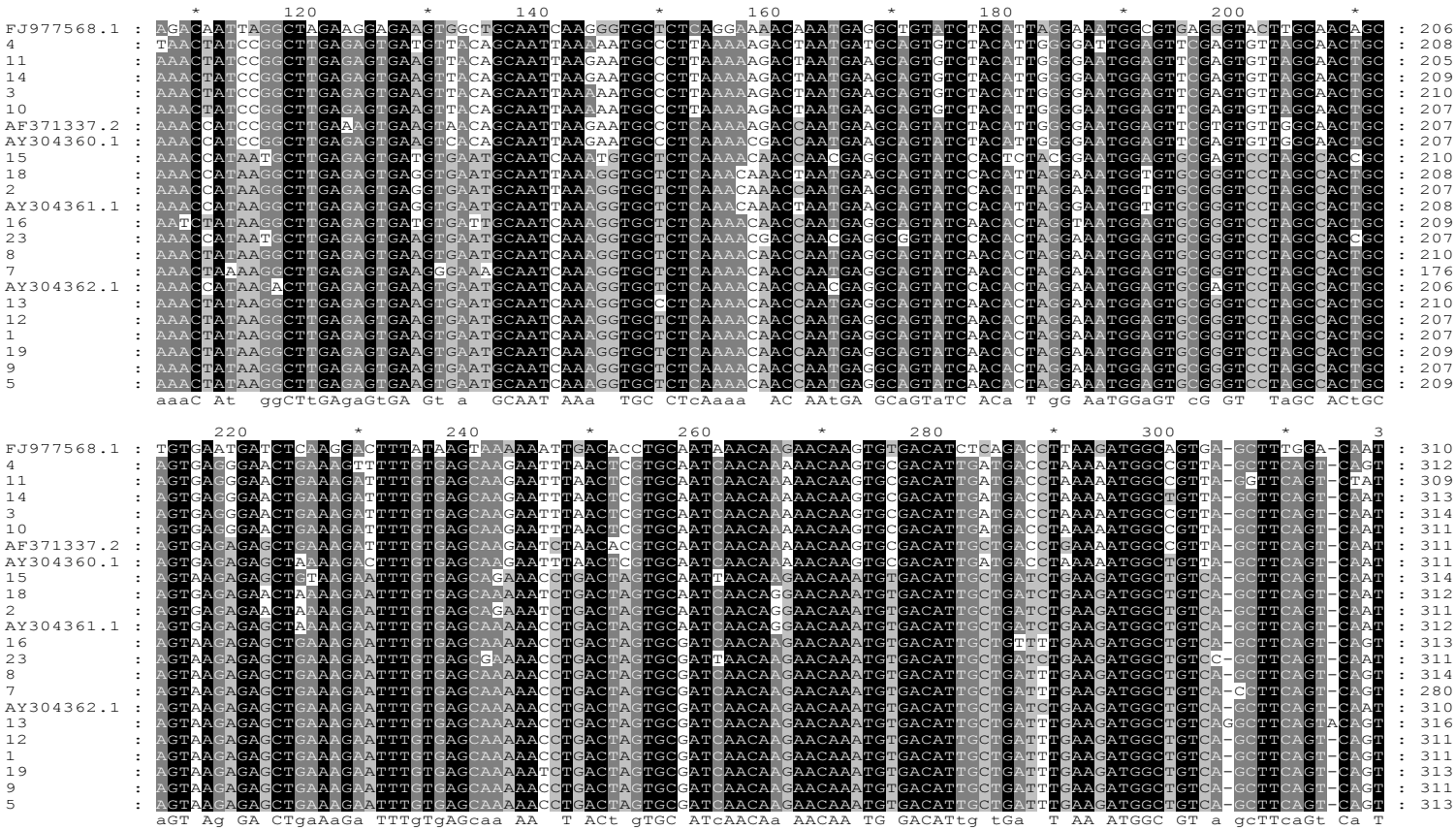


Cycle sequencing products which are negatively charged migrate by charge from the cathode (Negative) to the Anode (positive) through a polymer matrix according to termination length or molecular weight of each sequence i.e. the shortest fragment moves through fastest and the fluorescence given from the terminating ddNTP is read by the camera/fluorescent detector of the 3130xl genetic analyzer.

APPENDIX E: Demographic details of Lwak and Kibera study sites including hMPV frequencies.

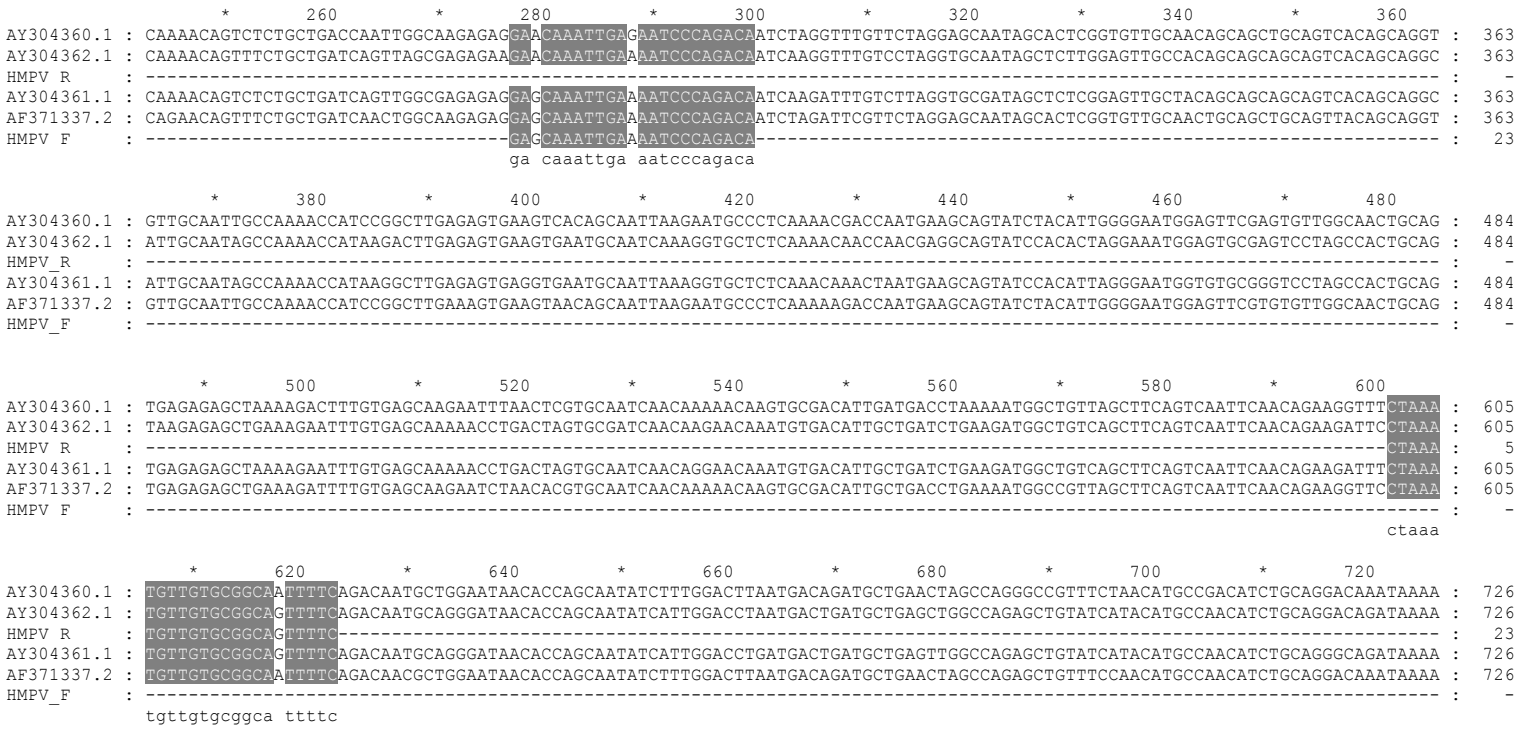
		LWAK					KIBERA				
Age Group	**Sex**	**Estimated Population under Study**	**NP/OP swabs done**	**Person years**	**hMPV +**	**incidence per 1000 person years**	**Estimated Population under Study**	**NP/OP swabs done**	**Person years**	**hMPV +**	**incidence per 1000 person years**
					TOTAL	TOTAL				TOTAL	TOTAL
<12 mnths	M	403	93	1,246.17	4	3.21	408	158	1186.11	9	7.59
	F	389	90	1,214.68	1	0.82	422	141	1,240.24	7	5.64
12-23 mnths	M	424	87	1,257.24	5	3.98	500	122	1920.37	10	5.21
	F	499	72	1,282.64	2	1.56	568	130	1,958.83	10	5.11
24 – 59 mnths	M	1,226	129	3,482.91	3	0.86	1,532	196	4581.37	11	2.4
	F	1,277	131	3,515.10	4	1.14	1,604	170	4,509.51	13	2.88
5 – 17 years	M	5,177	226	14,378.83	6	0.42	4,295	196	11380.9	11	0.97
	F	4,904	259	13,622.79	10	0.73	4,639	224	12,611.89	9	0.71
18 – 34 years	M	3,310	79	8,321.24	1	0.12	4,319	100	14531.9	7	0.48
	F	3,558	157	8,946.32	3	0.34	5,047	230	15,490.33	4	0.26
35 – 49 years	M	966	60	2,862.40	1	0.35	2,035	63	5942.08	1	0.17
	F	1,615	77	4,782.90	1	0.21	1,307	46	3,436.30	0	0
50+ years	M	1,354	45	3,953.40	0	0.00	601	14	2,260.93	0	0
	F	2,192	78	6,403.94	2	0.31	196	4	798.2	0	0
TOTAL	**M**	12,860	719	**36,265.36**	20	0.55	**13,690**	**849**	41,803.60	**49**	1.17
	F	14,434	864	**39,005.20**	23	0.59	**13,783**	**945**	40,045.29	**43**	1.07
< 5yrs	**M**	**2,053**	309	**5,988.57**	12	2.00	**2,440**	**476**	7,687.85	**30**	3.9
	F	**2,165**	293	**6,010.17**	7	1.16	**2,594**	**441**	7,708.58	**30**	3.89
Adults	**M**	**10,807**	410	**29,383.43**	8	0.27	**11,250**	**373**	4,115.76	**19**	0.56
	F	**12,269**	671	**33,888.39**	16	0.49	**11,189**	**504**	2,336.72	**13**	0.4

APPENDIX F: Multiple sequence alignment prior to phylogeny



Multiple sequence analysis showing sequences with reference to AY 304361.1 from position 103 to 312

APPENDIX G: Multiple sequence alignment showing binding sites for the primers on all the four hMPV subtypes



APPENDIX H: Nucleic acid extraction using Qiagen viral RNA minikit

- Thaw samples on ice, but work at room temperature (15-25^{0}C).
- Vortex the NP/OP sample and then add 100 µl of this to 400 µl aliquot of buffer AVL-carrier RNA solution.
- Mix thoroughly by vortexing for 15 seconds so as to ensure efficient lysis. Incubate at room temperature for 10 minutes to allow for complete lysis of the virus.
- Add 400 µl of ethanol (96-100%) to each sample and mix thoroughly by vortexing for 15 seconds.
- Transfer 630µl of each sample from the previous step into the column supplied in the kit. Centrifuge the column (1 minute at 8000 rpm at room temperature) and discard the filtrate.
- Place the column in a clean 2ml collection tube. Transfer the rest of the sample into the appropriate column. Centrifuge the column (1 minute at 8000 rpm at room temperature) and discard the filtrate.
- Place the column in a clean 2ml collection tube, add 500µl of AW1 wash buffer to each column, centrifuge the column (1 minute at 8000 rpm at room temperature) and discard the filtrate. Place in a clean 2ml collection tube.
- Add 500µl of AW2 wash buffer to each column and centrifuge (3 minutes at 14,000 rpm at room temperature) and discard the filtrate. Place in a clean 2ml collection tube.
- Centrifuge again without adding anything (1 minute at 14,000 rpm at room temperature).
- Place in a clean 1.5ml microcentrifuge tube, and then add 100µl AVE elution buffer equilibrated to room temperature to columns and incubate for 10 minutes at room temperature. Centrifuge for 1 minute at 10,000 rpm at room temperature and store the eluted viral RNA at -80^{0}C for long term storage.