MOLECULAR DIAGNOSIS OF COINFECTIONS WITH HUMAN BOCAVIRUS AND GENOTYPING KENYAN ISOLATES OF THE EMERGING VIRUS

Misigo Dennis Mwala (BSc)

P.O. Box 2116-00202 KNH, Nairobi, Kenya

Email: dennismisigo@yahoo.co.uk

This project report is submitted in fulfilment for the award of the degree of Master of Science in Tropical and Infectious Diseases (MSc TID)

University of Nairobi Institute of Tropical and Infectious Diseases (UNITID)

NOV 24TH, 2010

Declaration

This research project is my original work and has not been presented for academic award in any university.

Nov 24th, 2010

Misigo Dennis Mwala

This research project has been submitted for the award of the degree of Master of Science in Tropical and Infectious Diseases with approval as University supervisors for Dennis Misigo Mwala

25/11/10

Dr David N. Mburu

Institute of Tropical and Infectious Diseases

University of Nairobi

25/11/2010

Dr Dufton M. Mwaengo

Department of Medical Microbiology

University of Nairobi

Abstract

In Kenya, most surveys on viral respiratory tract infections, often, have always concentrated on the classical pathogens, which are respiratory syncytial virus, adenovirus, rhinovirus, parainfluenza and influenza virus as the commonest viral causes of upper respiratory tract infections whereas newer pathogens have emerged that cause similar symptoms. Human bocavirus (HBoV) is one of the recently discovered viruses isolated from respiratory specimens globally. This virus may cause respiratory symptoms clinically indistinguishable from influenza. The primary aim of this study was to generate data that will provide baseline information on a newer and potential aetiological agent of acute respiratory tract infections in our communities. The data generated confirms that human bocavirus is already circulating in Kenya and can be isolated from clinical specimens of patients experiencing influenza like symptoms. Only 1.8 % of samples analyzed were positive for HBoV DNA, making the frequency of detection of HBoV less common than all the other classic pathogens but probably more common than herpes simplex virus type 1. Human bocavirus infection in Kenya seems to affect children less than two years of age indicating that infection with the virus might primarily be during the early stages of life. A BLAST search of GenBank with sequences obtained from Kenyan HBoV isolates identified several human bocaviruses that showed 99% identity based on VP1/2 region. Furthermore, phylogenetic analysis of the Kenyan isolates and several other previously sequenced viruses (HBoV1-4) indicated that the HBoV1-like lineage (ST2 genotype) is the main strain circulating in Kenya.There is need to understand more on the role of the virus in cases of lower respiratory tract and gastrointestinal infections, because patients with HBoV infections present with non-specific clinical manifestations. This is the first report documenting the existence of HBoV in Kenya and describing the genotype. This information should add to the body of knowledge needed to monitor for emerging viral infections.

Acknowledgment

I acknowledge the excellent help provided by my supervisors Dr Mwaengo and Dr Mburu during the drafting stages of the project and finally for editing the dissertation manuscript.

I am also grateful to Eric Magembe (PhD fellow at ILRI) for his helpful suggestions concerning the analysis of our data especially by training me on the use of basic bioinformatics tools.

I shall always be thankful to the International Livestock Research Institute staff at segolilab who assisted in the procurement of the primers and in particular Ben Kiawa and Lucy Muthui for the assistance with the sequencing.

Abbreviations

µl	Microlitre
AAV	Adeno- Associated Viruses
ABI	Applied Biosytems
ADV	Adenovirus
AMDV	Alleuinia Mink Disease Virus
bp	Base Pair
BPV	Bovine Parvovirus
TBE	Tris Borate EDTA buffer
CI	Confidence Interval
CnMV	Canine Minute Virus
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotide Triphosphates
ELISA	Enzyme Linked Immunosorbent Assay
EM	Electron Microscope
ENT	Enteroviruses
ERC	Ethical Review Committee
HBoV	Human Bocavirus
HCOV	Human Coronaviruses
HMPV	Human Metapneumovirus
HSV	Herpes Simplex Virus
IFA	Immunofluorescence Assay
Kb	Kilo Base Pair
KDHS	Kenya Demographic and Health Survey
KNH	Kenyatta National Hospital
MPHS	Ministry of Public Health and Sanitation
mV	Millivolts
nm	Nanometre
NP	Nucleoprotein
NS	Non Structural Protein
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
rpm	Rotations per minute
RSV	Respiratory Syncytial Virus
SARS	Severe Acute Respiratory Syndrome
SPSS	Statistical Package for the Social Sciences
ST	Stockholm Strain Prototype
TTV	Torque Teno Virus
UON	University of Nairobi
VLPS	Viral- Like Particles
VP	Viral Capsid Protein
XMRV	Xenotropic Murine Leukaemia Virus
^{o}C	Degree centrigrade

Contents

List of figures

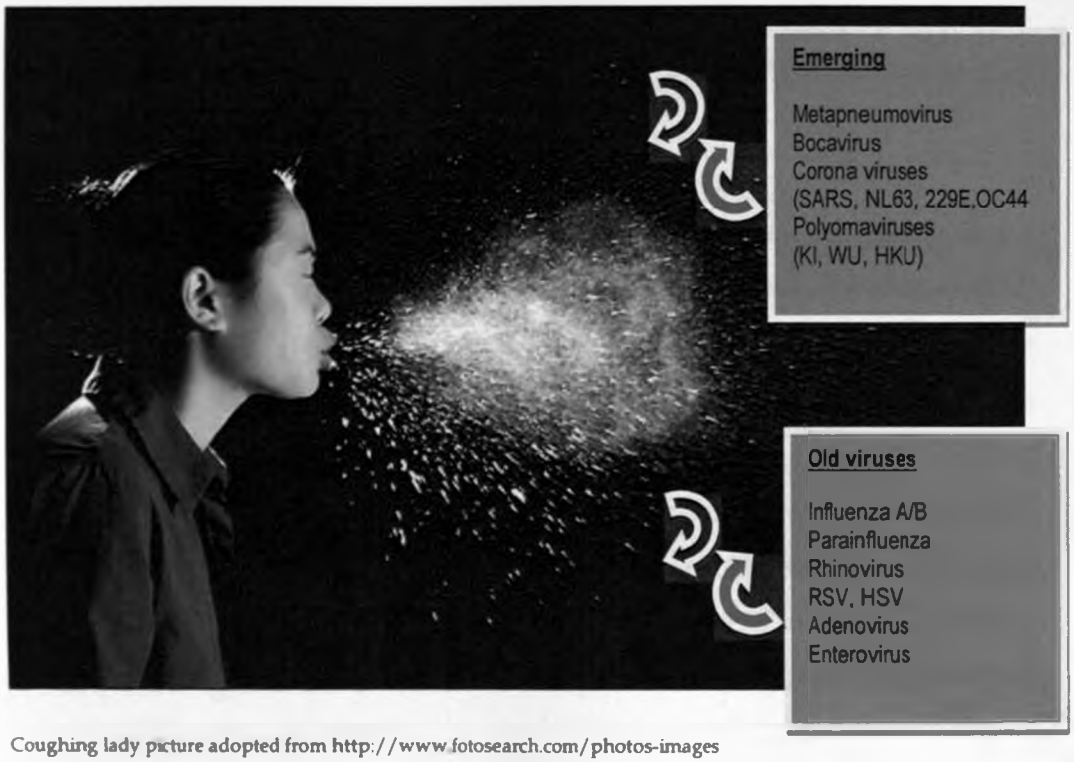
1.0 CHAPTER ONE

1.1 Introduction

1.1.1 Emerging respiratory agents

In the past fifteen years several previously undescribed viruses have been identified from humans. Among these were the blood borne viruses, hepatitis G in 1995 (Simons et al., 1995) and the TT virus in autoimmune disorders (Nishizawa et al., 1997). Thereafter, several viruses were reported such as XMRV in prostate cancer (Urisman et al., 2006), merkel cell polyomavirus in skin cancer (Feng et al., 2008) and others through analysis of clinical specimens from the human respiratory tract. Among viruses associated with the respiratory tract discovered during this period (**Figure 1**) included metapneumovirus (HMPV) in 2001(van den Hoogen et al., 2001), three new human coronaviruses (HCoV); the severe acute respiratory syndrome (SARS) associated coronavirus in 2003 (Marra et al., 2003), Coronavirus NL63 (NL63) in 2004 (van der Hoek et al., 2004), Coronavirus HKU1 (HKU1) in 2005 (Woo et al., 2005), human bocavirus (HBoV) in 2005 (Allander et al., 2005), polyomaviruses KI (KIV) and WU (WUV) in 2007 (Allander et al., 2007a) and the recently described human rhinovirus type C (Lau et al., 2009). It is the advances in diagnostic procedures that have lead to these discoveries, but even with the most sensitive molecular techniques, only 40-60% of infections of upper and lower respiratory tract have been associated consistently with a causative microorganisms.

FIGURE 1. The number of new viruses that have been discovered in addition to the classical causes of upper respiratory tract infections



Coughing lady picture adopted from http://www.fotosearch.com/photos-images

The factors contributing to the emergence and resurgence of infectious disease are multiple and complex. They include demographic changes, changing lifestyles, population growth, political instability, global travel and trade (Luna et al., 2007), development of new technologies, ecological changes, microbial adaptation or mutations and deterioration of national infrastructures for the control of infectious agents (Morse, 1995). Some variants of classical pathogens such as the pandemic H1NI influenza and avian influenza although not all that new are providing challenges in a environment where old diseases are potentially being caused by either totally new agents, mutants, recombinants or simply some of them are crossing species barriers through zoonosis with and without an ability to adapt to human-human transmission.

The genera bocavirus comprises only two type species found mainly in the upper respiratory and the gut of animals, majority of which are host specific. Human bocavirus (HBoV) is a putative member of the family parvoviridae and the genus bocavirus, a classification based on the genomic similarity to the bovine parvovirus and canine minute virus. Until recently human parvovirus B19 had been the only virus in this family with demonstrated pathogenicity for humans. Since the first report and identification of human bocavirus in 2005 and eventual designation of the virus into two genotypes, many investigators have confirmed the presence and prevalence of HBoV worldwide (Allander, 2008).

The genome of HBoV has been fully sequenced but a lot is still unknown about its transmission, pathogenicity and distribution. There is also incomplete proof of association of HBoV with disease. However, various studies have associated HBoV as a co-infection with other pathogenic bacteria and viruses (Arden et al., 2006) responsible for respiratory tract infections. It may be assumed that the simultaneous presence of more than one viral pathogen would be associated with epidemiological and clinical features that differ from single infections, especially with respect to disease type and severity.

Human bocavirus was variably detected in 1.5%–19% of individuals with acute respiratory illness, with a noted frequency of detection of between 5.0–5.5 % for most of the studies. Most of the studies have found that the virus is most frequently detected in infants less than 5 years of age. However, infections are not just limited to children but adults are infected as well (Kupfer et al., 2006; Longtin et al., 2008). While most studies have detected the virus during the winter season other reports describe a constant occurrence throughout all seasons.

Kenya has approximately 39 million people (Kenya demographic health survey report of 2009) with urban centers that experience significant international traffic. It is possible that HBoV has already been introduced into Kenya and can be isolated from clinical specimens of patients experiencing influenza like symptoms. The frequency of detection makes HBoV less common than respiratory syncytial virus and influenza viruses but approximately as common as, human metapneumovirus, parainfluenza virus type 3, and adenoviruses; and probably more common than coronaviruses.

This project attempts to detect such an event through selected archived supernatants of samples collected over a two and half year period between January 2007 and June 2009. We envisaged that the diagnosis of HBoV would form a useful adjunct to focused influenza-like illness surveillance when capacity for this diagnosis is available.

1.1.2 Justification

It is now well established that most respiratory infections are typically disease of household's, schools and institutions, which characteristically represent situations where individuals are herded together at susceptible ages. Within our boundaries, it is estimated that country wide, 56% of all cases of outpatient clinic consultations are reportedly due to repeated episodes of acute respiratory tract infections in children below five years of age (Kenya Demographic Health Survey Report 2009). These children generally present with complains of cough accompanied by congestion and difficulty in breathing or fever. An assessment of the impact of this infections indicate that economically and socially the number of lost hours in terms of hospital visits and lost school days due to respiratory tract infections calls for intensified studies on aetiological agents responsible so that proper management of cases can be prescribed. This is well exemplified by the recently pandemic HINI influenza outbreak.

Analysis of data from different countries highlights difference in the importance of certain pathogens both bacteria and viruses. For instance, the rest of Africa with exception of South Africa is lagging behind in documenting of incidences of emerging respiratory pathogens. In Kenya, most surveys often have concentrated on the classical pathogens namely RSV, adenovirus rhinovirus, parainfluenza and influenza virus as the commonest viral causes of upper respiratory tract infection whereas newer pathogens have emerged that cause similar symptoms.

HBoV is one of the recently discovered viruses isolated from respiratory specimens globally. This virus may cause respiratory symptoms clinically indistinguishable from influenza.

An understanding of the natural history of HBoV is essential to determination of its incidence and prevalence in a community, but for us to understand it we must first isolate it and then determine its potential as a pathogen in our communities. The optimal samples to utilise are those samples from which other viral agents have already been isolated from because the virus is most frequently recovered in the presence of other underlying viral infections of the respiratory tract even though there is evidence that the virus can exist as a single infecting agent.

There is also need to describe the genetic similarities and divergences between isolates mainly because bocaviruses are highly diverse and are prone to recombination events (Kapoor et al., 2010b). To obtain this measure of genetic relatedness sequencing and typing of significant isolates has to be done. This will add value to the investigation by assisting in ascertaining potential commonalities of worldwide experiences with the viruses at the molecular level and hence exposure.

In summary, at the national level a collation of data about newly recognized or rare infectious agents is essential so that their epidemiology can be described and basis for future research is provided. In this study, we hoped to generate data that will provide baseline information on a newer and potential aetiological agent of acute respiratory tract infections in our communities. These will permit rational and strategic public health interventions to eliminate or control the hazard both in relations to the specific current problems of undiagnosed viruses of respiratory tract and the other potential for future incidents.

1.1.3 General objective

To isolate and determine the genotype of isolates of Human bocavirus identified in Kenya.

1.1.4 Specific objectives

1. To demonstrate presence of HBoV DNA in archived supernatants of samples collected for diagnosis of viruses causing influenza-like illness.
2. To sequence and genotype HBoV DNA based on the variable VP gene.
3. To submit for archiving the Kenyan HBoV genotype sequences to GenBank.

2.0 CHAPTER TWO

2.1 Epidemiology

2.1.1 Global incidence

It is now recognized that HBoV has a global distribution as incidences have been reported in Europe and America: Sweden (Allander et al., 2005), Germany (Weissbrich et al., 2006), the Netherlands (Monteny et al., 2007), France (Foulongne et al., 2006), Italy (Maggi et al., 2007), Switzerland (Regamey et al., 2007), United Kingdom (Manning et al., 2006), USA (Kesebir et al., 2006), Canada (Bastien et al., 2006), Australia (Sloots et al., 2006), Spain (Vila et al., 2008) and Israel (Hindiyeh et al., 2008). Information is also available in countries other than the UK and USA, for example in Asian and middle east represented by Jordan (Kaplan et al., 2006), China (Qu et al., 2007), Japan (Ma et al., 2006), Iran (Naghipour et al., 2007), Korea (Choi et al., 2006), Thailand (Fry et al., 2007). In this brief review it is not possible to list all but as noted by a review of databases in public domain extensive studies have not been carried out in Africa as evidence by reports only emanating from South Africa (Smuts and Hardie, 2006). Although the source of the present infection is not known it is difficult to assess the true prevalence of HBoV infections worldwide. It is probable that many cases occur in children but the sub-clinical nature of many viral illnesses together with limited diagnostic facilities explains why proportions on the virus are uncertain especially in Africa.

2.1.2 Age incidence

The incidence of HBoV is between 1.5% and 18% based on tests of respiratory samples from individuals with acute respiratory illness. The age incidence varies but is always highest in the preschool child. The infection is common during the first year of life and maximum incidence is reached by the second and third year of life. A high incidence occurs in infants under 2 years of age. The incidence is low in teenage children and young adults, but may increase in elderly men and women following a pattern similar to influenza infections. A number of host conditions predispose the individual to infectious with the virus and this includes patients having pre-existing medical conditions.

2.1.3 Manifestations

The carriage of HBoV in the absence of symptoms may occur during the incubation period, convalescent or in healthy persons. Patients convalescing from respiratory infections with virus or bacteria may harbour HBoV in the throat or nose for many weeks. In study of some patients, of those found positive, the clinical symptoms and disease manifestations observed in infected children include pneumonia, bronchiolitis, wheezing, respiratory distress, asthma (Vallet et al., 2009) hypoxia, fever, rhinitis, laryngeal croup and, more rarely, conjunctivitis or rashes (Lindner et al., 2008a, Simon et al., 2007). The virus has been isolated in otitis media cases (Beder et al., 2009). Otitis media primarily affects children and viral upper respiratory infections often are responsible for the events that lead to it. There are case reports of HBoV in the immunosuppressed due to T-cell deficiency (Kainulainen et al 2008, de Vries et al., 2009) and also dissemination in transplant patients (Fischer, 2008).

In HBoV infection, a few cases report severe disease, others are associated with disease of a less obvious nature, and may appear to be non-pathogenic not surprisingly, the pathogenic status is uncertain. The isolation of HBoV from healthy subjects does not prove that it merely a commensal; conversely isolation form disease tissue, particularity when other microorganisms are present, does not prove pathogenicity. However, because patients with HBoV infections present with non-specific clinical manifestations, laboratory testing is necessary to confirm a diagnosis.

2.1.4 Transmission and biological samples

The real distribution of HBoV cannot be determined with certainty. HBoV-DNA has been detected in tonsils and adenoids (Clement et al., 2009), serum (Kantola et al., 2008), faecal (Kapoor et al., 2009), urine (Pozo et al., 2007) samples obtained from infants and water supply (Hamza et al., 2009). It is believed, however, that transmission may occur in close contact settings via respiratory droplets or by hand to mouth suggesting an oral route of entry with invasion through the gut. Other parvoviruses have been conveyed in donations of blood or tissue (Fryer et al., 2007a). HBoV survives in water for prolonged periods. For most respiratory infections the clinical sample can be collected either by swabbing or aspiration from the nasopharynx or oropharynx area.

2.1.5 Comorbidities

HBoV is more prevalent as a co-infection than a single viral infection. It has been associated with infectious conditions caused by rhinovirus, cytomegalovirus, respiratory syncytial virus, adenovirus, parainfluenza (Manning et al., 2006), *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Klebsiella pneumoniae* and non-infectious conditions (Smuts and Hardie 2006). HBoV and airway infection has not been established yet, but some controlled studies supporting this hypothesis have been published (Allander et al., 2007b; Fry et al., 2007; Kesebir et al., 2006; Maggi et al., 2007; Manning et al., 2006, Simon et al., 2007).

2.1.6 Pathogenesis

As reviewed by McIntosh (2006) we know essentially nothing substantial about mode of transmission, the frequency of transmission, the year-to-year variation in frequency, the viremia, duration of viral shedding, the mechanism of recovery, the frequency of reinfection, the mechanism of immunity and the full spectrum of associated disease and complications. However, a number of viral disease states lead to the induction of cytokines, which alter a variety of cellular enzymatic paths and functions. Physiologically, cytokines cause nausea, vomiting, diarrhoea and fever. Phospholipases can stimulate cytokines. A phospholipase A2 motif has been identified on the VP1- unique region of all genomes of HBoV (Qu et al., 2008; Kapoor et al., 2010b). In parvoviruses this molecule is important for infectivity. Physiologically, phospholipase A2 hydrolyses phospholipid membranes releasing metabolites leading to the production of toxic products through two main pathways, the lipooxygenase and cyclooxygenase pathways.

The toxic leukotriennes generated through the lipooxygenase pathway can induce strong brochoconstriction and this might explain why HBoV infections exacerbate asthma attacks (Vallet et al., 2009) and also cause wheezing. HBoV infections also induce production of Th1 and Th2 cytokines also resulting in the mechanism responsible for wheezing (Chung et al., 2008)

2.1.7 Immunity to HBoV

It is apparent that up to 96% of health adults have past immunity to HBoV (Soderlund-vernermo et al., 2009). This is because human serum contains antibodies produced as result of exposure to HBoV this also suggests that HBoV is endemic. Elisa technique using viral particles for the VP gene has developed to the level of being used for routine evaluation of seroprevalence of antibodies to HBoV by showing a rising titre of immunoglobin G or presence of Immunoglobin M (Soderlund-vernermo et al., 2009). In a study by Lindner (2008b) HBoV specific IgM antibodies were detected predominantly in infants with HBoV viremia suggesting an association of HBoV infection with acute respiratory disease. Elisa will provide cheaper tools for future use in basic clinical laboratories. Because patients with HBoV infections present with non- specific clinical manifestations, laboratory testing is necessary to confirm diagnosis.

Several techniques have been evaluated, but serological testing remains the preferred approach to diagnosis because of its relative ease and simplicity. Newer cultivation techniques may provide acceptable alternatives. However, as observed for most respiratory infections neither serology nor isolations allows diagnosis to be made early enough to affect patient management.

2.2 Classification

By phylogenetic analysis, two new genera, *Amdovirus* and *Bocavirus*, within the subfamily of *Parvovirinae* were recently created (ICTVdB - The Universal Virus. Database, version 4) the *Amdovirus* genus contains only one species, the Aleutian mink disease virus (AMDV), while the Bocavirus genus contains two species which are bovine parvovirus (BPV) and canine minute virus (CnMV) hence, the name *bo*vine and *ca*nine). All members are autonomously replicating viruses that were earlier classified in the *Parvovirus* genus.

2.2.1 Virion structure

The structure of HBoV resembles other parvoviruses. The virion consists of a non- enveloped isometric capsid with icosahedral symmetry (ICTVdB), hexagonally shaped and about 25nm in size (Brieu et al., 2007).

2.2.2 Human parvoviruses

The human B19 is erythrotropic virus that targets the red cell precursors in bone marrow. It is common in humans causing both asymptomatic and symptomatic infections commonly associated with erythema infectiosum, transient aplastic crisis, arthritis, foetal death and persisted anaemia in immunocompromised. In a recent study, two new parvoviruses, PARV4 and PARV5, were detected in serum of HCV-positive Intra Venous Drug Users (IVDUs), but not in non-IVDUs (Fryer et al., 2007b). Earlier on PARV4 and a related variant PARV5 had been isolated in pooled plasma used for the manufacture of plasma products (Fryer et al., 2006) however, little is known about the role of these viruses in human disease.

2.3 Genomics

Full-length genome analysis of the human bocavirus showed that this virus is closely related to the bovine parvovirus and canine minute virus, both of which are members of the genus *Bocavirus*, family *Parvoviridae*. The virus contains a single linear positive-sense and negative-sense single-stranded deoxyribonucleic acid. The complete genome is between 4000 to 6000 nucleotides long (ICTVdB - The Universal Virus Database, version 4). The genome of bocaviruses is slightly larger than that of AMDV and it has different sequences at the two genomic ends (Schwartz *et al.*, 2002). Unlike any other parvoviruses it encodes a nuclear phosphoprotein (NP1) with unknown function(s) (Schwartz *et al.*, 2002). The genome of parvovirus consists of two major ORFs encoding a non-structural protein (NS1) and at least two capsid proteins (VP1 and VP2), respectively (**Figure 2**). Moreover, HBoV also has a third middle ORF encoding a non-structural protein (NP1) of unknown function (Allander et al., 2005). There are two genotypes of HBoV circulating worldwide, genotypes ST1 and ST2, which differ by only 26 nucleotides across the entire genome, with most of the conserved amino acid changes located in the VP1/VP2 open reading frame (Kesebir et al., 2006). On the African continent the current documentation shows that the ST2 lineage circulates in South Africa while in Asia and Europe and the Americas both genotypes have been isolated.

2.3.1 Conserved and variable regions

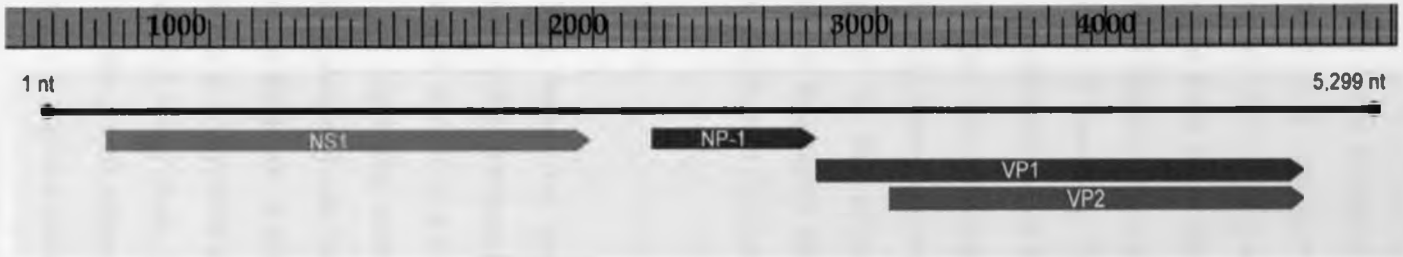
An evaluation of the different gene segments indicate that for the purposes of diagnosis the nucleoprotein 1 gene would be the preferred segment because it is more conserved and it is only found in members of bocavirus genus and not in any of the other parvoviridae. However, for the purpose of investigating the diversity of lineages, the gene for the viral capsid protein 1 and 2 is the most variable and highly informative region.

2.4 Animal bocaviruses

In animals, these viruses were demonstrated way back in the 1960's. Bocavirus has been isolated in dogs, cattle, and pigs and recently in gorillas (Kapoor et al., 2010a). The bocavirus minute virus of canines (MVC) was previously thought to be harmless until it was demonstrated as the cause for respiratory and enteric infections in puppies (Manteufel and Truyen (2008). The isolation of the virus in animals opens up some speculation on the role wild and other domesticated animals play in the epidemiology of HBoV infections. For instance, the pig plays an important role in reassortment and recombination of influenza virus as they act as the mixing vessels could the same be said of bocaviruses. As Zehender et al (2010) were evaluating the evolution of HBoV using Bayesian coalescent inference; they found a recombined isolate of HBoV that could not be regarded as belonging to either the ST1 or ST2 lineage. This is a pointer to importance of recombination events in this group of viruses.

FIGURE 2. The genome of human bocavirus.

The four genes of HBoV are labelled based on their presumed function, according to homologous genes in other parvoviruses. Non-structural protein 1 (NS1) is a DNA-binding protein involved in gene transcription. NP1 is also non-structural and is a highly conserved protein of unknown function. The capsid proteins are viral protein 1 (VP1) and viral protein 2 (VP2). The VP1 gene has phospholipase A2 motif, which important in tissue tropism and infectivity in other parvoviruses.



http://www.ncbi.nlm.nih.gov/nuccore?term=bocavirus

2.5 Bioinformatics

2.5.1 Multiple alignments

Constructing a multiple alignment corresponds to developing a hypothesis of how a number of sequences have evolved through the processes of character substitution, insertion and deletion. The input to multiple alignment algorithms is a number of homologous sequences i.e. sequences that share a common ancestor and most often also share molecular function. Often the first step in a chain of bioinformatical analyses is to construct a multiple alignment of a number of homolog DNA or protein sequences. Once a multiple alignment is constructed it can form the basis for a number of analyses. One is to investigate the phylogenetic relationship of the sequences by tree-building methods based on the alignment. It can be used to search for conserved regions in the alignment that can be prime candidates for holding functionally important sites. And lastly, comparative bioinformatical analysis can be performed to identify functionally important regions.

2.5.2 Distance based methods

Two common algorithms, both based on pairwise distances, are the Unweighted Pair Group Method using Arithmetic averages (UPGMA) and the Neighbor Joining algorithms. A simple but popular clustering algorithm for distance data is UPGMA. The algorithm assumes that the distance data has the so-called molecular clock property i.e. the divergences of sequences occur at the same constant rate at all parts of the tree. This means that the leaves of UPGMA trees all line up at the extant sequences and that a root is estimated as part of the procedure (Lemey et al., 2009). The neighbor-joining algorithm, on the other hand, builds a tree where the evolutionary rates are free to differ in different lineages, i.e., the tree does not have a particular root. The method works very much like UPGMA but with some differences in the way the nodes are generated. Other algorithms include Parsimony, Maximum likelihood and Bayesian methods the latter are probabilistic

methods of inference. Both have the pleasing properties of using explicit models of molecular evolution and allowing for rigorous statistical inference. However, both approaches are very computer intensive. The neighbour join algorithm is generally considered to be fairly good and is widely used (Lemey et al., 2009).

2.5.3 Inferring phylogenetic trees

In molecular epidemiology of infectious diseases, phylogenetic inference is important for a given set of aligned sequences and it is possible to infer their evolutionary relationships. The tree consists of a number of nodes (also termed vertices) and branches (also termed edges). These nodes can represent an individual, a species, or a higher grouping and are thus broadly termed taxonomical units. In this case, the terminal nodes (also called leaves or tips of the tree) represent extant species of a family and are the operational taxonomical units (OTUs) (Lemey et al., 2009). The internal nodes are termed hypothetical taxonomical units since they are not directly observable. The ordering of the nodes determines the tree topology and describes how lineages have diverged over the course of evolution. The branches of the tree represent the amount of evolutionary divergence between two nodes in the tree and can be based on different measurements. A tree is completely specified by its topology and the set of all edge lengths (Gregory, 2008).

2.5.4 Bootstrap values

A popular way of evaluating the reliability of an inferred phylogenetic tree is bootstrap analysis, which is also a measure of confidence in the generated branches. A high bootstrap score is a sign of greater reliability.

2.6 Diagnostic methods

2.6.1 Culture and animal models

HBoV is as unique as the several other newly discovered viruses in that it is difficult for all of these viruses to conform to Koch's postulate (Williams 2010), since neither of them can or has been cultured nor has an animal model of infection been established to study them (McIntosh, 2006). No known method is suitable for routine virus isolation yet meaning that detection usually relies exclusively on probing for the presence of the viral genome (Mahony, 2008).

2.6.2 Electron microscopy

The electron microscope is the most valuable technique in the study of virus particle structure but it is impractical for routine diagnostic. The presence of parvovirus-like particles in clinical samples found positive for HBoV DNA by PCR confirms that HBoV presents the structural characteristics of the *Parvoviridae* family members (Brieu et al., 2007).

2.6.3 Elisa diagnosis

Diagnostic laboratory assays are useful to distinguish HBoV infections from other causes of viral respiratory infections, to determine the presence of protective antibodies or recent exposure. In one study the major capsid gene VP2 of HBoV was recombinantly expressed in insect cells. The Expression of VP2 in insect cells led to the formation of virus-like particles (VLPs), which had the typical icosahedral appearance of parvoviruses. Purified VLPs were then used to establish an ELISA for detection of IgG antibodies against HBoV in human sera (Lin et al., 2008). The drawback is that Elisa tests for Immunoglobulin G is useful for seroprevalence studies but is not helpful in diagnosing acute infection because it simply indicates past exposure to virus. Fourfold or greater rises in titre remains the mainstay for confirming diagnosis.

Enzyme linked immunosorbent assay as a diagnostic technique provides objective end point and allows rapid testing of many individual specimens, however HBoV antigen for use in Elisa at the moment require recombinant technology making it extremely expensive to produce and not readily available. In the near future serological testing may still remain the preferred approved diagnosis procedure because of its relative ease and simplicity.

2.6.4 Hemagluttination

Lin et al (2008) discovered that the VLPs of HBoV did not hemagglutinate human red blood cells even at a high concentration, whereas purified BPV had hemagglutinated. The result from the hemagglutination assay suggested that HBoV might be more like the human adeno-associated virus type 2 (AAV2) in that they do not hemagglutinate human red blood cells. The inability to be detected by hemagglutination assays is representative of the difficulties in detecting or isolating this virus from clinical specimens.

2.6.5 Novel molecular methods

New technologies as reviewed such as VIDISCA (Van der Hoek et al., 2004), pan-viral DNA microarrays and high throughput sequencing have stimulated efforts to identify novel viruses in the respiratory tract and in other human disease states (Sloots et al., 2008). The discovery of human bocavirus is one good example of how random PCR-cloning sequencing is being applied to identify potential pathogens (Allander et al., 2005). In resource limited situations, where sequence based genotyping is not necessary, the use of restriction length polymorphism assays on PCR products (Ditt, Viazov et al. 2008) would provide a cheaper alternative.

2.6.6 Polymerase chain reaction

Recent advances in molecular biology, particularly the introduction of the polymerase chain reaction (PCR) assay, have greatly improved the detection of viral respiratory pathogens (Sloots et al., 2008). A variety of primers have been designed for conventional and real time PCR (Kleines et al., 2007) based on the VP1/2, NS-1 and NP-1 gene sequences all showing great sensitivity and specificity for HBoV (Smuts et al., 2008). Chieochansin et al (2007) showed that the NS1 and NP1 genes are the most conserved regions and thus, will not demonstrate differences between HBoV isolates. As it has become apparent from the phylogenetic tree of the complete coding sequence, variations of HBoV are shown in VP1 and VP2 based on which the virus can be divided into two and three groups, respectively (Smuts et al., 2006). These findings indicated that although HBoV is a conserved virus, the sequence analysis showed that amino acid variations almost uniformly appear in the capsid protein (VP1 and VP2). The NS1 and NP1 represent the most conserved regions, which should be utilized for HBoV detection.

Most of the epidemiological data that has been generated recently in many studies are based on experiments using the NSI, NP1 and VP1/2 gene for HBoV detection. Nested PCR is a variation of the polymerase chain reaction (PCR), in that two pairs (instead of one pair) of PCR primers are used to amplify a fragment. The two sets of primers amplify a specific DNA fragment using two separate runs of PCR. The second pair of primers fragment (as they lie or are nested within the first) function to amplify a smaller specific DNA fragment located within the first PCR product. The advantage of nested PCR is that if the wrong PCR fragment was amplified in the initial run, the probability is quite low that the region would be amplified a second time by the second set of primers. Thus, Nested PCR is a very specific PCR amplification (Knipe et al., 2007).

2.7 Prevention, control and treatment

The fact that HBoV is often co-detected with other respiratory viruses and that more needs to be understood about the mode of transmission, distribution, persistence, severity of infections or co–infections and immunity. It is still difficult to draw conclusions about effective control measures and chemotherapeutic agents against the virus (Schildgen et al., 2008).

Neske and colleagues (2007) found that 50% of HBoV stool was from children with HBoV positive nasal aspirates and that the viral load was high in this group than in a comparison group of HBoV negative stool that lead them to hypothesize that HBoV is swallowed during respiratory tract infection and subsequently excreted in stool without it replicating in the gut. However, the other members of the genus (GBV, MVC and BPV) cause diarrhoea in gorillas, dogs and cattle.

Hamza et al (2009) findings on river water with HBoV DNA indicates that just like the other members of its family that this virus is stable in the environment and is also suggestive of contamination with an infected humans excreta because the same virus has a role in causing gastroenteritis just like rotavirus and adenovirus though the strains may be different. The implications of this observation are that in most resource scarce setting poor hygiene and sanitary conditions are good source for virus transmission. The quality of drinking water can be compromised at any point between the original sources and the ultimate consumption. However, the classical methods of selecting the best available source of water and then providing the barriers of storage, filtration and disinfections could be effective in preventing transmissions provided that the distribution systems is also protected and maintained.

The challenge with all respiratory viruses is that neither serology nor isolation allows a diagnosis of most causative agents to be made early enough to affect patient management. Even though the diagnosis of HBoV will not result in specific antiviral therapy it will reduce and eliminate the excessive use of antibacterials to treat respiratory infections. Further studies will need to be developed to enable determine the stability and viability of the virus in the presence of common disinfectants. This information will only be obtainable when efficient *in vitro* culture systems or experiments on animal models are established.

3.0 CHAPTER THREE

3.1 Design and methodology

3.1.1 Study population

a. Inclusion criteria

The specimens selected for assaying were all archived supernatants that were harvested from cultures within 48hrs after primary inoculation. The supernatants were collected from specimens that were positive for herpes simplex virus type 1 (HSV), respiratory syncytial virus (RSV), parainfluenza virus (Para), adenoviruses (Adv) and enteroviruses (Ent), which represent nasopharyngeal swabs collected between January 2007 and June 2009. The positive specimens were confirmed by Immunofluorescence assays using virus- specific antibodies for each genus.

b. Exclusion criteria

All influenza positive supernatants and supernatants for samples collected before January 2007 and after June 2009.

3.1.2 Samples size

Our study design was within the limits of a descriptive retrospective cross-sectional study where we expect a binary outcome. To calculate the samples size the formulae $Z^2\ p\hat{}\ (1-p\hat{})/M^2$ was employed. Where $Z=1.96$, $p\hat{}= 0.5$ and $M=0.05$.

We considered a confidence level at 95% and within a confidence interval (M) of 5%, the estimated proportion (p^) at 50% because the true prevalence is unknown worldwide. This generated a figure 384.16 [$(1.96)^2$ x 0.5x 0.5 / $(0.05)^2$]; hence we randomly select and analyzed 384 supernatants. Approximately 915 specimens had been confirmed to be positive by culture method for other respiratory viruses. We selected only 384 required samples from these 915 positive samples aided by a random integer generator (http://www.random.org/integers/).

3.2 Study regions

Selection of representative specimens for this study was based on the samples collected during an influenza survey, which represent the major district hospitals and three capital cities, Nairobi, Kisumu, Malindi, Busia, Mombassa, Isiolo, Kisii and Kericho, that experiences frequent international movement **(Figure 3).**

3.3 Materials and Methods

3.3.1 Specimens storage

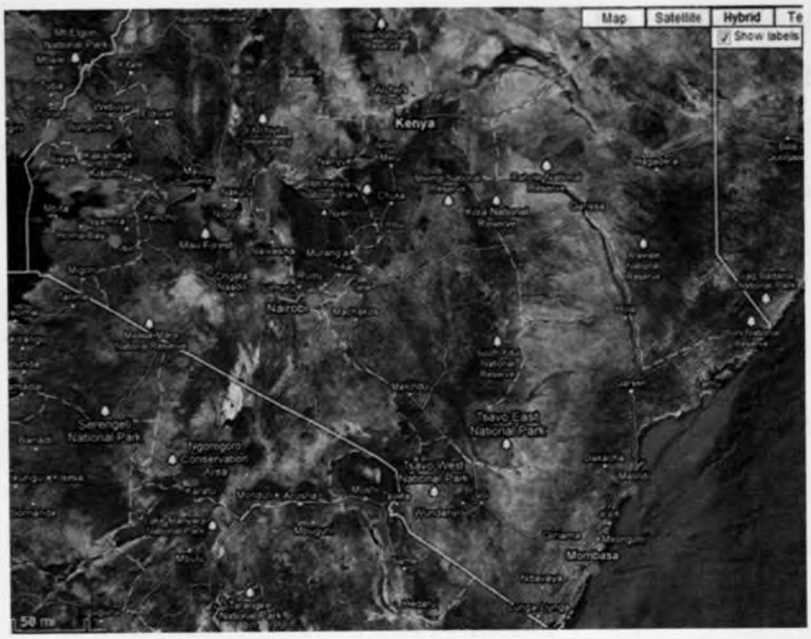
The supernatants (stored in 1.0 ml mixture of viral transport media and culture media containing Dubelco's modified eagles medium supplemented (DMEM) with 2 % fetal bovine serum and 1% antibiotics in a cryovial) were retrieved from ultra low temperatures (- 80^0c) freezers and rapidly thawed at 37^0C. The supernatants were reassigned with new unique project numbers. However, they were linked to the original samples for more information during data analysis.

3.3.2 DNA extraction

DNA was extracted from 50-100µl of the samples using the Qiagen QIAmp DNA blood minikit (Qiagen, Germany) procedures. Because of the limited resources and assumed rare exposure to the virus the initial screening involved pooling of 25 µl of each samples in batches of 10 to make ninety two pools each with total volume of 250 µl prior to DNA extraction. The positives pools were then set aside and individual's samples in each positive pool analyzed.

The QIAamp DNA purification kit (Qiagen, Germany) was used to extract DNA as per the manufacturer's recommendation. In summary, 250µl of sample was lysed with solution AL .The lysed solution was mixed with Qiagen Protease (DNase/ RNase free) and incubated at 56°C for 10 mins. The lysate was precipitated with 99% ethanol and loaded onto the QIAamp Mini spin column. The DNA was adsorbed onto the QIAamp silica membrane by centrifugation. DNA bound to the QIAamp membrane was then washed in two centrifugation steps at 8000 rpm and 14,000 rpm respectively by the use of two different wash buffers, Buffer AW1 and Buffer AW2 respectively. The Wash conditions ensured complete removal of any residual contaminants. Purified DNA was eluted from the QIAamp Mini spin column in a concentrated form in either water (60µl) for immediate sequencing or Buffer AE. Yields were increased by incubating with the elution solutions at room temperature for 5 minutes before centrifugation at 14,000rpm. If the purified DNA was to be stored, elution in Buffer AE (10 mM Tris·Cl; 0.5 mM EDTA; pH 9.0) was preferred and the extracts stored at -20°C because DNA stored in water is subject to degradation by acid hydrolysis.

FIGURE 3. A Google map of Kenya with eight sites marked in either apple green or red



The apple green marked sites are regions from which the specimens screened were positive for HBoV DNA. (Satellite map adopted from www.googlemap.com)

3.3.3 Nested PCR

The virus has remained undetected until recently because of their inability to replicate *in vitro* and this makes molecular diagnostics as the only best method to detect the viruses. The main PCR methodology was a semi-nested PCR similar to that employed by Smuts et al (2006). The details of the primer sequences are in Table 1. We used external and internal primers designed for the NP1, VP1/VP2 genes. We expect to amplify a 368bp product of NP1 gene and confirm it by amplifying a 980bp product of the VP1/VP2 gene respectively. The VP2 sequence is nested within VP1 gene and the NP1 gene is absent from other parvoviruses but it is present in bocavirus (**Figure 2**).

3.3.4 Primer sequences

Table 1. Oligonucleotide primers for PCR amplification and sequencing of HBoV targeting the viral capsid protein (VP1/2) and nucleoprotein (NP1).

Gene	Forward	Reverse
VP1/2 outer	**VP-A** (5'GCA CTT CTG TAT CAG ATG CCT T-3)	**VP-B** (5'CGT GGT ATG TAG GCG TGT AG-3')
VP1/2 inner	**VP-C** (5'CTT AGA ACT GGT GAG AGC ACT G-3')	**VP-B** (5'CGT GGT ATG TAG GCG TGT AG-3')
NP-1 outer	**NP-A** (5'TAA CTG CTC CAG CAA GTC CTC CA-3')	**NP-B** (5'-GAA GCT CTG TGT TGA CTG AAT-3')
NP-1 inner	**NP-C** (5'CTC ACC TGC GAG CTC TGT AAG TA-3')	**NP-B** (5'-GAA GCT CTG TGT TGAC TGA AT-3')

Adapted from Smuts and Hardie (2006)

3.3.5 DNA amplification

A 50 µl PCR reaction was set up containing 1X PCR buffer, 0.2 mM of each primer (VP-A and VP-B, NP-A and NP-B), 0.2 mM of dNTP's mix, 1.5 mM $MgCl_2$, 2 Units of Taq polymerase and 10µl of extracted DNA in 50µl aqueous reactions. Cycling profile include an initial denaturing at 94°C for 10 minutes, this was followed by 40 cycles of denaturing at 94° C (30sec), annealing at 55°C (30 sec) and 72°C (1 min). There was a final extension at 72°C (5 mins) after the 40 cycles and holding at 4°C. Two microlitre of the primary PCR product was used as a template in the nested reaction with primers VP-B and VP-C, NP-B and NP-C in separate reaction tubes in the secondary PCR under the same cycling conditions

3.3.6 Gel electrophoresis

Ten microlitres of amplicons were loaded on gels and electrophoresis was performed using 1.5% agarose containing 2µl of ethidium bromide. The gels were run at 100mV-150mV in 1X Tris Borate EDTA buffer pH 8.0. Two markers, a standard 100-bp and 1Kb DNA ladder, were used as a reference. The evidence of amplified DNA was ascertained from bands that could be visualized at the specified and expected sizes and photographed on a UVP trans-illuminator installed with the alpha Imager ™ software (alpha innotech).

3.3.7 Clean up-and purification of DNA

The amplicons for sequencing were purified using the Qiagen QIAquick PCR purification kit from gels or solutions depending on the level of purity (Qiagen, Germany). Briefly, one volume of amplified PCR product was mixed with five volumes of Buffer PB and then adsorbed onto silica-gel a membrane column, which has selective binding properties. Salt residues were quantitatively washed away by the ethanol containing Buffer PE. Any residual Buffer PE, which could have interfered with subsequent enzymatic reactions, was removed by an additional centrifugation step at 14,000 rpm for one minute. The bound DNA was eventually eluted after a minute incubation at room temperature with 30 µl of Buffer EB (10 mM Tris·Cl, pH 8.5), or water. The eluted DNA was stored at -20°C until required for further analysis.

3.3.8 Amplicons concentration and purity

The concentration of DNA was determined spectrophotometrically using a Thermo Scientific NanoDrop™ 2000/2000c spectrophotometer on 0.5-2 µl of the samples against blanked water controls. The sample intensities along with the blank intensities were used to calculate the sample absorbance using the Beer-Lambert equation. The concentration of the DNA was determined at 260nm .The ratio of absorbance at 260 and 280 nm was used to assess the purity of DNA. A ratio of ~1.8 and above was generally accepted as "pure" for DNA. If the ratio was appreciably lower in either case, it indicated the presence of protein or other contaminants that absorb strongly at or near 280 nm and thus had to be re-purified.

3.4 Sequencing

3.4.1 Cycle sequencing

Nucleotide sequencing was carried out on the sense and anti-sense strands of HBoV DNA using the ABI Prism BigDye ™ Terminator cycle ready reaction kit v 2.0 as recommended by the manufacturer. The reaction was in 20 µl reactions. Briefly, 50-100ng of PCR product of the VP1/2 gene was mixed with 4 µl of terminator ready reaction mix (A, C, G and T-dye terminators; dNTPs, Tris HCl pH9.0, $MgCl_2$, thermal stable pyrophosphatase and amplitag DNA polymers), 2 ul of 5x sequencing buffer and 3.2 pmol primer. Cycle sequence was performed according to the protocol specified for the Gene Amp 9600 themocycler (Applied Biosytems, Germany), which is outlined as preheating at 94^{0}C for 1min, strand separation at 94 ^{0}C for 10 secs, annealing at 50^{0}C for 5 seconds, and extension at 60 ^{0}C for 4 mins for 25 cycles and finally held at 4^{0}C infinitely. The primers used for sequencing were similar to those used for the internal round of amplification of the VP gene (Primer VP-B and VP-C).

3.4.2 Sequence product purification

Excess deoxy terminator was removed from the DNA sequencing products with centrim - Sep ™ columns. The purified products were dried in a vacuum centrifuge for 10–15 minutes at -60 ^{0}C and stored at -20^{0}C.

3.4.3 Automated sequencing

Automated sequencing was performed on an ABI Prism 3100 Genetic analyzer version 2 (Applied Biosytems, Germany). The raw data sequence was performed with sequencer software version 4.0.5.

3.4.4 Raw data editing

Initially we performed a blind BLAST at the NCBI website to screen for similarity of our nucleotides sequences in comparison to those in records in the GenBank database (Benson et al., 2010). The raw ABI sequencing data was analyzed to check for contaminating sequences, using the DNA baser v 2.91.5 sequence assembly software (Heracle Software, Germany). We were able to automatically clean the sequences by trimming off the low quality bases at the end of the chromatogram, assembled all the sequences into a contig, and corrected base pair errors such as deletions. The Fasta output from the single and multiple contig was then ready to be used for alignment. The contiguous consensus sequences used to generate the phylogenetic tree are shown in the **appendix.**

3.4.5 Phylogenetic analysis

The Kenyan isolates were compared to known reference sequences of bocavirus (HBoV 1, HBoV 2, HBoV 3, HBoV 4 and parvovirus isolates (canine minute virus and bovine parvovirus) and few other selected sequences of the virus isolates from South Africa (ZA190/04, ZA2591/04), China (HK5), Europe (Bonn-1) and America (CRD2). These comparison sequences had accession numbers DQ317556.1 (ZA 2591/04), Bovine parvovirus (NC_001540.1), Canine minute virus (NC_004442.1), EF450721.1 HK5, NC_014358.1 (Bocavirus gorilla), NC_007455.1 (Human bocavirus 1), NC_012042.1 (Human bocavirus 2), NC_012564.1 (Human bocavirus 3), NC_012729.2 (Human bocavirus 4), FJ858259.1 (Bonn-1) and (DQ340570.1). (CRD2). The sequences for comparisons were downloaded as Fasta files from the GenBank database and aligned together with five of our sequences labelled KE-4S9/04, KE-9N7/11, KE-10N8/3, KE-48G7/8 and KE-23A8/11 in ClustalW using the multiple alignment algorithms. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4.1 (Tamura et al., 2006). The Phylograms were constructed with Mega 4 version beta 3 using kimura two-distance-based parameters via neighbour-joining method generated from 1000 bootstrap replicates. The codon positions included were 1st+2nd+3rd. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons.

3.5 Data analysis

The data analyses focused on determining HBoV as a viral agent of respiratory infections among diverse geographical and demographic groups. In particular we looked at associations between laboratory results and data from questionnaires where the age, sex, residence history including recent history of influenza like illness were be ascertained retrospectively from the clinical parameters documented in the database and cross-tabulation of proportions and frequencies analyzed by SPSS version 13.0 (SPSS Inc., USA) and EPI info™ (CDC, Atlanta)

3.6 Nucleotide sequence submissions

The sequences data were uploaded and directly submitted via the sequin software to NCBI. The five sequences of HBoV-VP1/VP2 gene representing DNA of five samples were deposited in GenBank and have been assigned with accession numbers HQ288041 (KE-9N7/11), HQ288042 (KE-23A8/11), HQ288043 (KE-10N8/3), HQ326235 (KE- 48G7/8) and HQ326236 (KE-4S9/04) respectively. The cleaned contig sequences submitted were between 733 – 790bp and are shown in the **appendix.**

3.7 Clearance

We got the ethical clearance from the KNH-ERC to proceed with the study under the ERC number P85/3/2010.

4.0 CHAPTER FOUR

4.1 RESULTS

4.1.1 HBoV isolates

HBoV DNA was amplified from seven samples out of 384-screened specimens (1.8%). This was in five females and two males. Retrospective reviews of records for age revealed that two (28.5%) of the seven were less than 12 months of age (8months precisely) and the other five were between one and 2 years of age (71.5%). The number of HBoV isolates per collection sites was four from Kisumu (57%) and one each from Busia, Nairobi and Kisii, respectively. No HBoV was isolate from the specimens representing other sites. Among the HBoV positive samples the coinfections with a single virus was observed as two with parainfluenza 3 only (2.5%), two with adenovirus only (2%) and one with enterovirus only (1.4%). Multiple infections consisting of HBoV plus two other viruses, which collectively accounted for 3% was observed when enterovirus was isolated together with adenovirus and also separately with herpes simplex virus type 1. This is summarised in Table 2 and **Figure 4**.

4.1.2 Coinfecting agents

From the 384 specimens screened the proportion of viral agents identified from the supernatants excluding HBoV was as follows; adenovirus only (26.5%), parainfluenza (20%), RSV (16.5%), enterovirus (18.1%), HSV (1.2%) and dual or multiple infections as a combination of any one of these viruses accounted for the remaining percentage (**Figure 5**).

4.1.3 Site distribution

The distribution of total specimens screened by geographical location was that 22.1% were from Kisumu, Nairobi (17.5%), Kisii (17.3%), Kericho (13.3%), Malindi (10.4%), Mombasa (9.9%), Busia (5.2%) and 4.2% were from Isiolo. In general, the number of samples analysed from all the sites based on the months of collection seemed to be relatively evenly distributed.

4.1.4 Age and sex distribution

A retrospective analysis of the age distribution of the patients from whom the original samples had been taken indicated that half of all the specimens assayed in this study (51.3%) were obtained from children between 1-2yrs, 29 percent were children below one year, and 18.5% were subjects older than 3years but less than 7years. 0.9% were adult specimens. There was a slight predominance of males (56%) in comparison to the females at 44%. This was statistically not significant ($p > 0.05$).

4.1.5 Month of isolation

In a stratified analysis of samples based on the month in which the swab was collected showed that the virus was isolated in all the three years represented by the samples screened. Specifically, four of the isolates were in 2007 specimens, these Isolates of 2007 occurred in the months of January, March, August and November, while the two isolated in 2008 samples indicated that the infections had occurred in the months of March and November. The 2009 isolates were in the month of April only. Hence HBoV was isolated twice, consistently in November and March, in two different years and from two different regions (Table 2)

Table 2. A summary of clinical and demographic data on the seven individuals positive for HBoV DNA

HBOV POSITIVE SUBJECTS (n=7)							
Demographics							
Subject No	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Age in months	22	22	13	8	13	14	8
Sex	M	M	F	F	F	F	F
Geo-location	Busia	Nairobi	Kisumu	Kisumu	Kisumu	Kisumu	Kisii
Period of isolation	Nov 08	Aug 07	Jan 07	Oct 07	Nov 07	Mar 08	Apr 09
Clinical Presentation							
Temp ≥38°C	√	√	√	√	√	√	√
Cough, nasal stuffiness, runny nose	√	√	√	√	√	√	√
Breathing difficulty	√	√	-	-	-	-	-
Diarrhoea	-	-	-	-	-	-	-
Vomiting	-	-	√	√	√	√	√
Fatigue, malaise, poor appetite	√	-	-	-	-	-	-
Neurological- restlessness	-	√	-	-	-	-	-
Coinfections	Para 3	hAdv	Ent	hAdv	Ent/ hAdv	HSV/Ent	Para 3

Abbreviations: Para 3; parainfluenza 3, hAdv; human adenovirus, Ent; enteroviruses, HSV; herpes simplex virus; √ = symptom present, -; symptom absent, M; Male F; Female, Nov 07; November 2007, Aug 08; August 2008, Mar 08; March 2008, Apr 09; April 2009, Jan 07; January 2007, Oct 07; October 2007,Nov 08; November 2008 **1**= KE-23A8/11, **2**= KE-48G7/8, 3= KE-38N7/01, 4= KE-9N7/10, 5=9N7/11,6=KE-10N8/3, 7=4S9/04

FIGURE 4. The percentage of HBoV positives among the coinfections

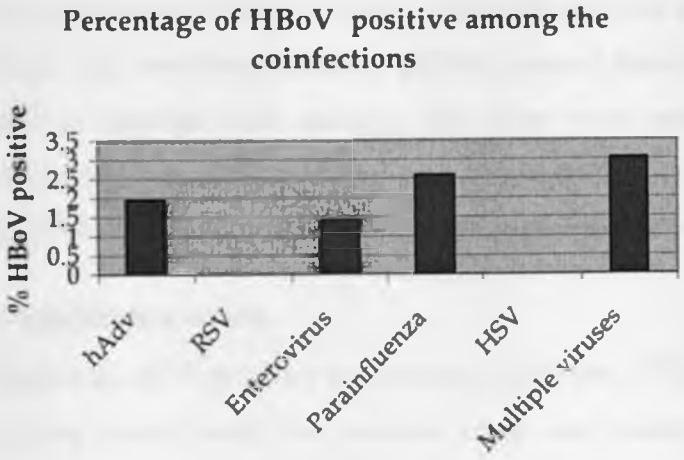
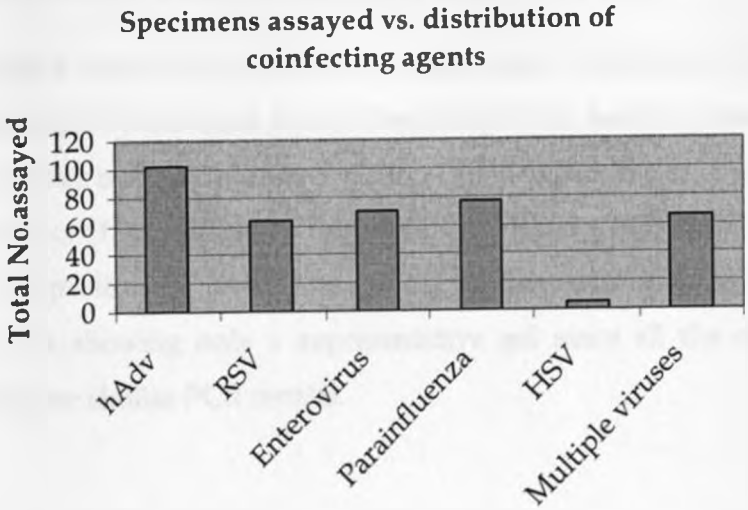


FIGURE 5. The distribution of virus coinfections



4.1.6 Clinical symptoms

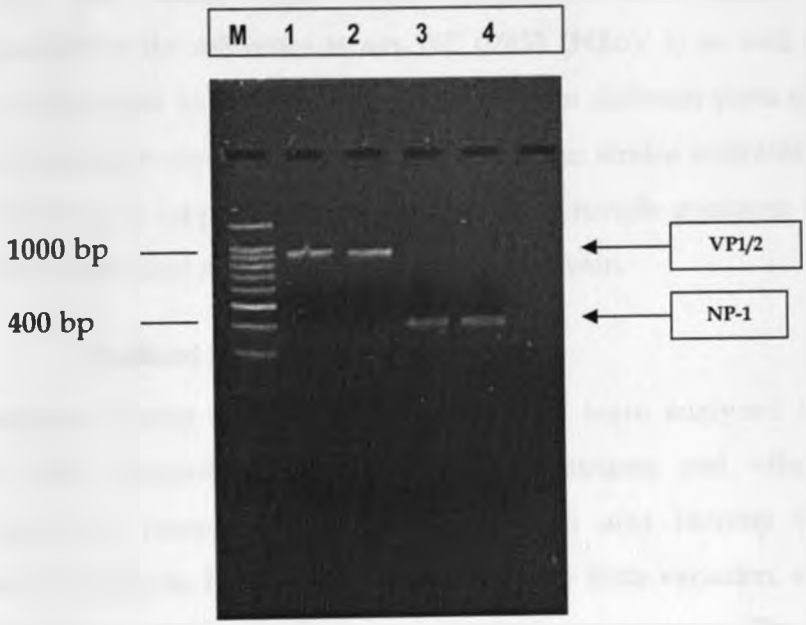
A retrospective analysis of the clinical parameters recorded in the archived questionnaires, as summarised in the table 2, revealed that all the seven individuals complained of fever, coughs, nasal stuffiness and runny noses. In addition, three had breathing difficulty (42.8%), one of them complained of fatigue, loss of appetite and malaise, the other was restless and had neurological symptoms and vomiting was observed in five of the seven individuals (71.4%). This was also statistically not significant ($p>0.05$)

4.2 PCR amplification results

Figure 4 shows the PCR product representing both the NP1 and the VP1/2 gene. The first round with the external (data not shown) NP primers amplified a 368pcr product that was confirmed by a second round PCR run with the internal primers (**Table 1**). For the purpose of sequencing the VP gene we amplified a final product of 980bp using both internal and external primers.. We successfully quantified and purified them on average they had a 260/280nm ratio of purity score range of about 1.7 to 1.83.

The samples were run on 1.5 % TBE agarose gel. The expected product size was estimated by the band size of the 100bp DNA ladder shown in Lane M (100bp ladder markers). Lanes 1 (KS9/11) and 2 (KE-10N811) show VP gene PCR product of 980bp, while Lanes 3 (KS9/11) and 4 (KE-10N811) show NP1 gene PCR product of 398bp.The gel electrophoresis results as illustrated in **Figure 6** is showing only a representative gel since all the other positive samples gave similar PCR results.

FIGURE 6. PCR products for both the VP and NP genes on agarose gel visualized as bands using ethidium bromide on a UV Trans-illuniminator



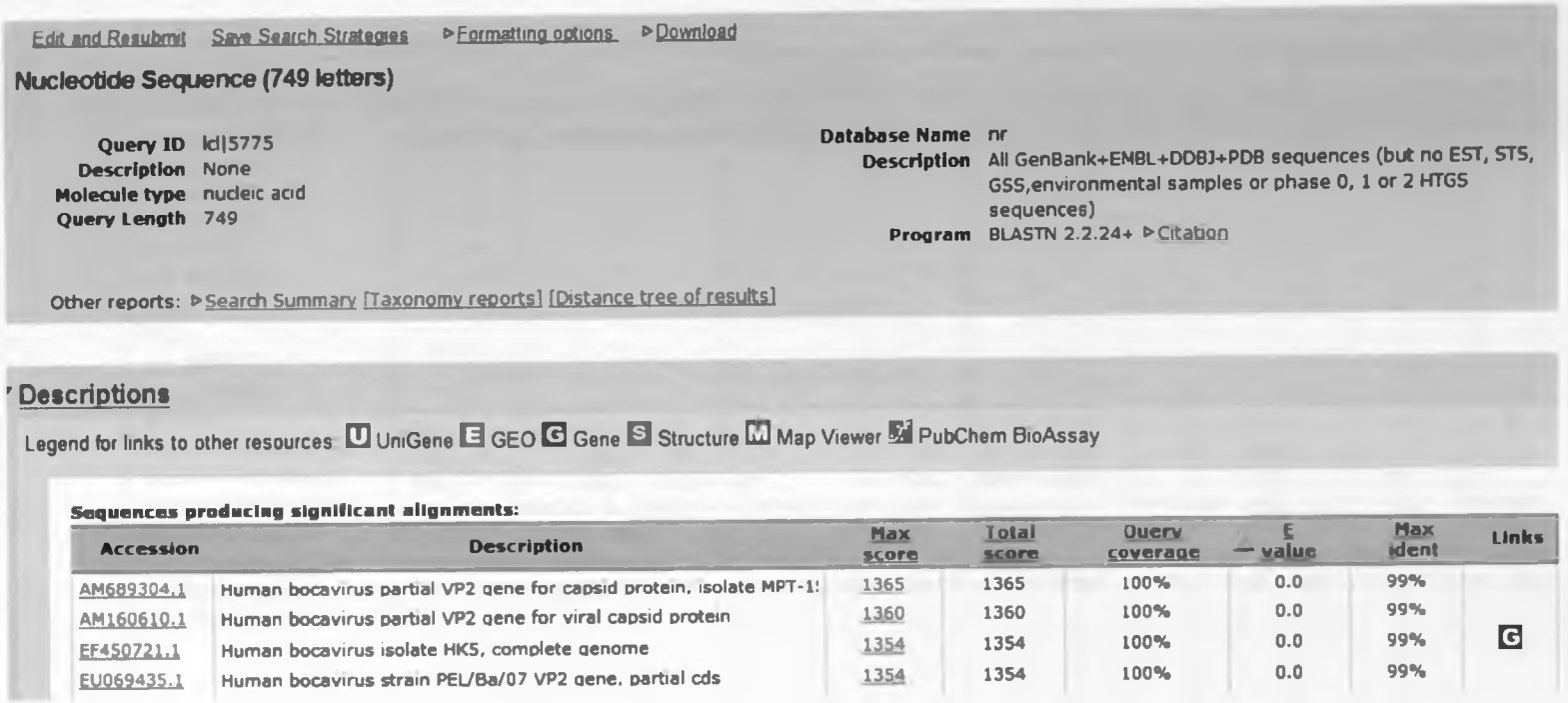
4.2.1 NCBI blast search

A search for homologies between our sequence and sequences in the public database indicated that our sequences had 99 % homology to human bocavirus meaning that they shared high homology with corresponding types of the isolates published in GenBank **(Figure 7).** In particular our isolates had similar sequences and amino acid arrangements to the reference strain, NC 07455 (HBoV 1) as well as to other strains closer to this reference isolated from different parts of the world. Nucleotide sequence analysis of the Kenyan strains indicated that the VP1/2 region amplified corresponded to nucleotide positions 4102-5002 of the published sequence of the prototype strain.

4.2.2 Predicted amino acid sequences

The deduced amino acid sequences alignment were analyzed in Bioedit software by comparing to the out group viruses and other human bocaviruses. A comparative analysis of amino acid identity **(Figure 8)** between the Kenyan HBoV isolates showed very little variation, and where differences were observed they were mainly conservative. The predicted peptide amino acids also had little divergence from those of HBoV 1. The greatest difference was between our isolates and the out group virus and the other HBoV viruses. This data further supports the observation that our isolates are closer to HBoV 1 species than any other HBoV species.

FIGURE 7. Example of a BLAST search result generated at NCBI web pages showing sequences producing significant alignments to one of our isolate in relation to the database records of other HBoV sequences.



Edit and Resubmit | Save Search Strategies | ▷Formatting options | ▷Download

Nucleotide Sequence (749 letters)

Query ID lcl|5775
Description None
Molecule type nucleic acid
Query Length 749

Database Name nr
Description All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS,environmental samples or phase 0, 1 or 2 HTGS sequences)
Program BLASTN 2.2.24+ ▷Citation

Other reports: ▷Search Summary [Taxonomy reports] [Distance tree of results]

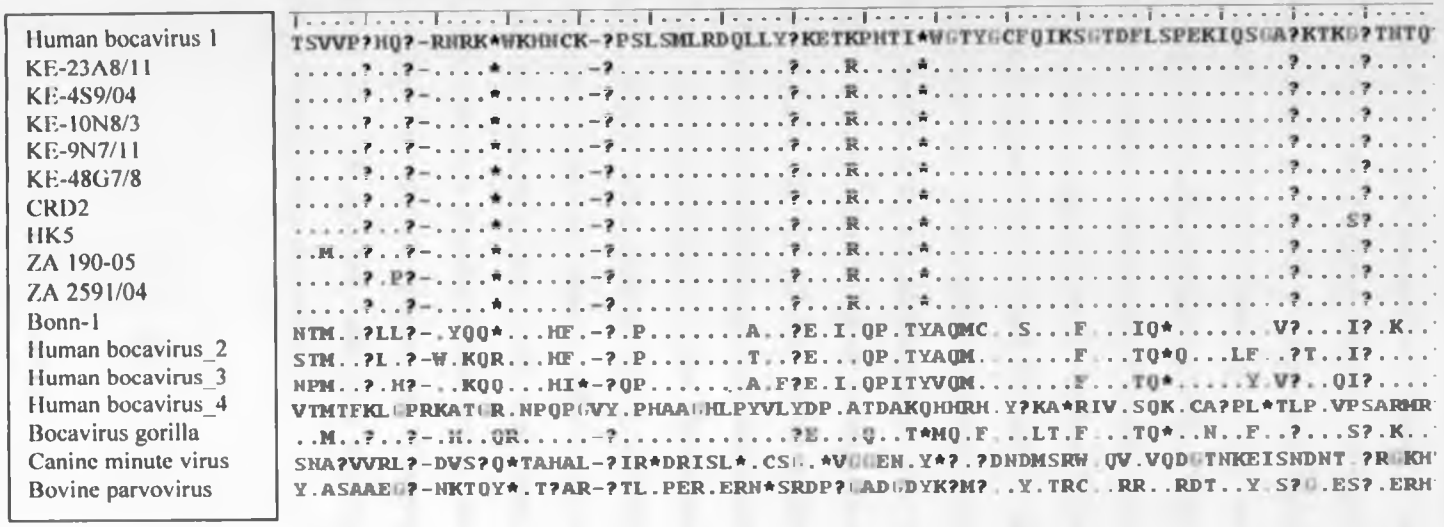
Descriptions

Legend for links to other resources: U UniGene E GEO G Gene S Structure M Map Viewer PubChem BioAssay

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
AM689304.1	Human bocavirus partial VP2 gene for capsid protein, isolate MPT-1!	1365	1365	100%	0.0	99%	
AM160610.1	Human bocavirus partial VP2 gene for viral capsid protein	1360	1360	100%	0.0	99%	
EF450721.1	Human bocavirus isolate HK5, complete genome	1354	1354	100%	0.0	99%	G
EU069435.1	Human bocavirus strain PEL/Ba/07 VP2 gene, partial cds	1354	1354	100%	0.0	99%	

FIGURE 8. Comparison of a portion of the deduced amino acid alignments plotted identities and similarity of the Kenyan isolates against HBoV 1 as a standard

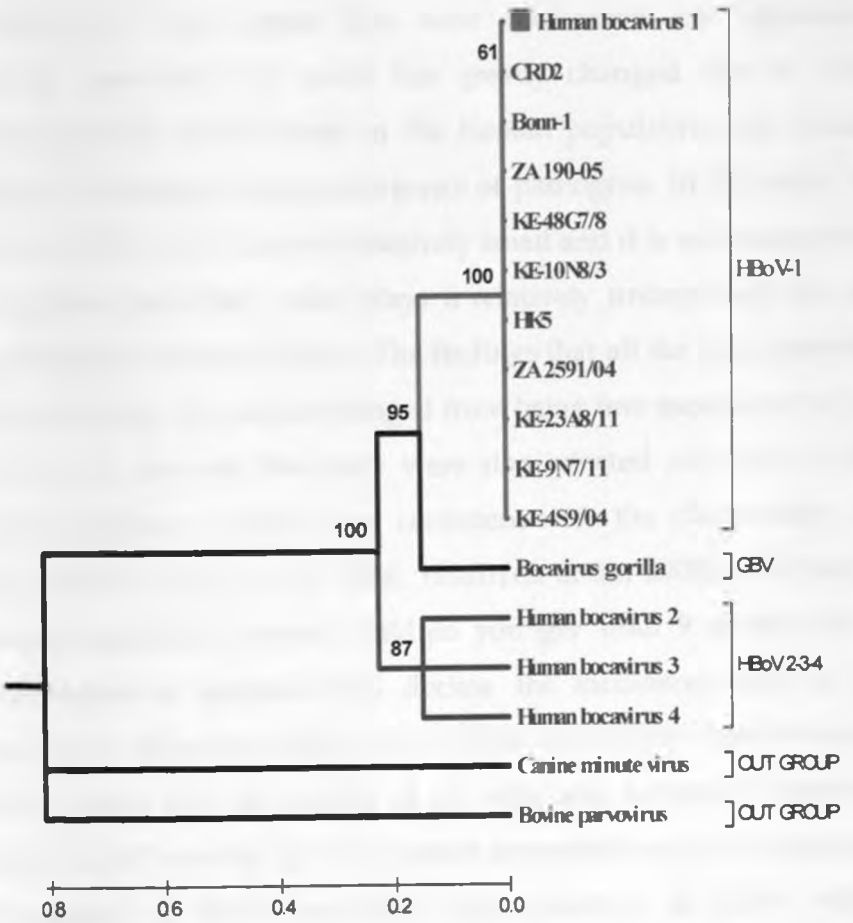


The dots represent the identical and similar amino acids and only the non- identical amino acids are indicated with alphabetical letters in different colours.

4.3 Phylogenetic tree clustering

Phylogenetic tree data is presented in Figure 9 and it shows very high bootstrap levels ranging from 61-100%. The canine minute virus and bovine parvoviruses, which are the out group viruses distinctively separated from the rest with a perfect bootstrap value of 100%. The other isolates were separated at three clusters. At the first level of separation, human bocaviruses 2, 3 and 4 separated from the rest with a high bootstrap value of 87% (blue branch). At the second level of separation the gorilla bocavirus is also separated with a high bootstrap of 95 % from the rest of the isolates, which also contains those isolated from Kenya. Interestingly, when compared to the human bocavirus 2, 3 and 4 species (blue branch) human virus 1 (red branch) clusters closely with those from Kenya than any of the other human bocaviruses. In general four clades were observed on the phylogenetic tree with the Kenyan strains clustering closer to the South Africa isolates as well as to Bonn-1, HK5 and CRD2 **(Figure 8)**. Six of the seven isolates were available for sequencing but we had insufficient amount of some of the samples hence its data has not be included in the illustrations of this report. This evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. We got the same tree topology using the maximum likehood algorithm (data not shown).

FIGURE 9. Phylogenetic analysis and comparison of human and animal bocaviruses sequences of the VP1/2 gene region.



The analysis included five isolates originating from Kenya. Those obtained from this study have names starting with "KE"; those retrieved from public databases are labelled as HK5, CRD2, ZA, and Bonn-1. The reference DNA sequences for bovine parvovirus and canine minute virus that served as the out group sequences are indicated with the black branch below HBoV 2-3-4 on the tree.

5.0 CHAPTER FIVE

5.1 Discussion

5.1.1 The findings

Investigations of the role of viruses in causing acute respiratory infections in Kenya dates back to the mid 1980s when Hazlett et al. (1988) found a large proportion of viral agents that were considered less important then. However, currently that trend has greatly changed due to a complex interaction with other factors in the human population and environment resulting in resurgence and emergence of pathogens. In this study the total number of HBoV isolates was relatively small and it is unclear as to whether this implies that HBoV virus plays a relatively unimportant role in acute respiratory infections in Kenya. The findings that all the seven individuals in whom the virus was isolated ranged from being few months old to less than two years of age and that they were also infected with one of the other regular respiratory viruses was consistent with the observations of many other authors (Arden et al., 2006; Hindiyeh et al., 2008). It is thought that maternal antibodies protect children younger than 9 months but as the concentration of maternal IgG decline the incidences tend to increase. Karalar and colleagues (2010) observed that acute HBoV infections are rare in children older than 30 months of age who also happen to display higher levels of HBoV-specific IgG. It therefore seems that age is the most important determinant of HBoV infections. The incidence of HBoV infection is highlighted in the first 2years of life and diminishes progressively to a low level in older children and adults. In regard to the clinical presentation of infections involving HBoV, It seems that the general symptoms that usher in influenza like-illnesses are common to a wide variety of viral infections and thus there is no one single sign that can be said to be pathognomonic for HBoV infections.

The presence of vomiting in some of the individuals suggested gastric discomfort even though there was no record of diarrhoea episodes. These symptoms of gastroenteritis have been highlighted in other studies. For example, in a study in Italy, one of the reasons for seeking healthcare services for acute respiratory infection in 10% of the children in whom HBoV was isolated was because they had gastroenteritis (Esposito et al., 2008). However, inferring that HBoV was the cause of the vomiting in the individuals whose samples we analysed in the current study was not within the scope of this project considering that our approach was within the limits of a cross-sectional study. To make such inferences one will be required to use either cohort or case-control study designs.

5.1.2 Seasonal distribution

Kenya's different topographical regions experience distinct climates and the samples we analysed represented two seasons in two different years. Generally, we noted that the months in which the virus was isolated corresponds with the cold months and the rainfall season, except for the January isolate, especially in the Lake Basin area, which has two rainy seasons: the "long rains" extending roughly from March to June, and the "short rains" lasting from approximately October to December. The highlands of western Kenya have a single rainy season, lasting from March to September. Most of the isolates obtained were from the samples emanating from the western and Nyanza region of Kenya. Investigating for seasonality was not within our objective but we could assume that more infections with HBoV are likely to coincide with trends and epidemiological pattern of adenovirus and parainfluenza virus infections.

5.2 Validation

The methods described in this study were able to amplify and sequence HBoV DNA from archived supernatants. One disadvantage though is that molecular diagnostics require one to use some expensive instruments and reagents. In our study it was necessary to employ molecular techniques considering the unavailability of alternative less expensive diagnostic procedures. Additional drawbacks of molecular assays include the inability to distinguish between viable and non-viable virus or between actual disease and latent infection, which is necessary to classify the clinical conditions. This particular drawback on use of PCR method for diagnosis, greatly affects the attempts to show causation making the association concept predominate because it has been shown that HBoV is shed for long period (Blessing et al., 2009). Analytically, one critical disadvantage is the potential for cross contamination of samples causing false-positive results.

In our study we employed a nested PCR combined with sequencing of amplified products to ensure that the results obtained were both valid and reliable. The amplicons provided virus sequence information that enabled phylogenetic comparison of the viruses from clinical specimens to other, previously recognized and characterized strains of the virus, and thus it helped with the determination of the "molecular epidemiology" of the virus. Samples were sent for confirmatory diagnosis and sequencing at International Livestock Research Institute laboratories.

5.3 Whole genomic sequencing

The sequencing and genotyping was important because of the generally observed behaviour of viruses. Geographic and host range isolation is thought to be responsible for speciation in some virus families (van Regenmortel et al., 2007). The investigation in this study was limited to sequencing of a segment of the VP gene. We hypothesize that more HBoV genotypes are likely to emerge because of their potential to recombine. It will be important in the subsequent studies for full viral genomic sequencing to be undertaken in order to capture more details such as recombination events from our isolates.

5.4 ICTVdB and NCBI resources

This project relied heavily on finding an appropriate taxonomic position for our unknown isolates based on comparative sequence analysis. The complete inventory of viruses at International Committee on Taxonomy of Viruses database (ICTVdB), which is a well-established viral database, was a primary resource for information about biological properties of the parvoviridae family and genus bocavirus. However, human bocavirus is still not listed in the ICTV report as member of the genus bocaviruses at least by the time this report was being prepared. The exploration of genomic and protein databanks via the National Center Biotechnology Information (NCBI) collection of nucleotides website through the Basic Local Alignment Search Tool (BLAST) searches against a selected set of viral sequences played a major role in viral taxonomic classification. Through this resource we were able to confirm that our sequences were homologous to those of other bocaviruses.

5.5 Genetic lineages reclassification

All our isolates clustered together and within the same clade as the reference strain NC 07455, (HBoV1) which also has the GenBank accession number DQ00496 making it the original designated prototype of the ST2 genotype. Over the recent past newer isolates that show variance in the nucleotide sequences to the prototype strain have emerged leading to the division of HBoV genotypes in 4 species. These are HBoV 1, HBoV 2, HBoV 3 and HBoV 4. Kapoor et al (2010b) propose that future classification of new species of this virus should be based on HBoV strains showing >8% protein and >10% nucleotide difference in the complete VP1 gene, whereas those showing >1.5% protein and >5% nucleotide difference in the genome should be considered different genotypes. The ST2 genotype is synonymous to HBoV 1. Both HBoV 1 and HBoV 2 have been isolated in both respiratory tract and the gut but the other two seem more of enteric pathogens than respiratory. HBoV 3 is a result of recombination of HBoV 1 and HBoV 2. This property has also been observed in HBoV 4. Kapoor et al (2010b) observed that HBoV 4 could be as result of recombination between HBoV 2 and HBoV 3. The pattern of recombination is described to involve the junction between the NP1 and VP1 gene in that one virus genome can have the NS1/NP1 gene sequences of one genotype and the VP1/2 gene sequences of another genotype. This potential to recombine is also a property witnessed in animal parvoviruses (Shackelton et al., (2007).

5.6 Prospects of cell culture systems

Attempts to isolate HBoV from clinical respiratory material on currently available conventional cell lines, like LLC-MK2, HEp-2, Vero, and MRC-5 cells are reportedly not successful. In this study supernatants collected from culture media that had been exposed to clinical specimens were analyzed. Therefore, even though it was possible to isolate the virus from these specimens this still does not in any way mean that HBoV was able to grow in either of this cell cultures. What it implies is that the media harvested had traces of infected epithelial cells making it possible to detect HBoV DNA. However, the virus was recently cultured in pseudostratified human airway epithelium cell culture(Dijkman, Koekkoek et al. 2009), which is a system that resembles the human airways in morphology and function. This presents to future research activities a convenient tool to dissect the viral multiplication cycle, an opportunity to understand the pathogenesis of the virus, and a means to design chemotherapeutic agents or vaccines if need be. While HBoV can't be cultured on conventional media one of the animal bocavirus, canine minute virus type 1 can be cultured on a cell line that has never been distributed outside Europe, the WRCC/3878D cells, perhaps efforts should be made to evaluate if this cell line and others can support its growth.

5.7 Limitations

The first restriction was inadequate samples. The samples used had already been used for other assays and the remaining volumes were insufficient to carry out a battery of other assays. We also excluded all samples positive for influenza virus either as single infections or as confecting virus. There was limited information in record on other clinical symptoms and the samples represented only upper respiratory tract infections cases recruited. There is need to understand the situation in both cases of lower and upper respiratory tract infections, of particular interest would be the prevalence in lower respiratory tract because the lower respiratory infections are always more severe and usually demand immediate attention. We did not screen nasopharyngeal samples from patients however the culture media harvested from cultures had traces of patient samples. A study involving direct patient specimens should be carried out.

In relation to the study design, an investigation excluding all other viruses and demonstrating the presence of HBoV DNA in clinically ill patients in the absence of other viruses possible through retrospective serological investigations combined with cultures and viral load assays would build up strong evidence on the pathogenicity of HBoV to humans. Wide world attempts are being made towards such inference though none has used the proposed model of investigation. In our opinion, ideal sampling technique to elucidate the importance of the virus would be to collect three specimens for simultaneous screening for HBoV meaning that from the same study subject three specimens should be collected; blood for viremia, cytokines and viral load, stool and nasopharyngeal aspirate or swab for cultures then establish the role of HBoV in infections. The three samples would provide information on virus entry portal, predilections site, distribution mechanism and possible exit point. The sampling population should include outpatient, hospitalised and healthy individuals of all age groups.

5.8 Conclusion

The main objective of this report was to demonstrate the presence of HBoV in Kenya and describe the genotype in circulation. We have show that human bocavirus infection exists in Kenya essentially in children less than two years of age meaning that infection with the virus might primarily be during the early stages of life though this may not be the general rule. The HBoV 1 like lineage is the main strain circulating in Kenya. Through the years, parvoviruses have been of long interest to veterinarians than to the medical practitioners but the discovery of human bocavirus and other viruses has renewed the interest. It is changing how parvoviruses are perceived clinically even though considerable controversy exists concerning the exact role of parvovirus in respiratory tract infection. In our opinion, the information gathered should first of all greatly assist the Kenya Ministry of public health and sanitation in fulfilling their mission to monitor for emerging pathogens and potentially epidemic diseases.

Secondly, to surveillance systems and clinical services providers within the country, the study efforts have provided supplemental information on other possible causative agents of upper respiratory tract infections other than the principal and classical viral targets. We have reservations on recommending for routine diagnosis because first the mechanisms of pathogenesis have not been well described and importantly, most clinical microbiology laboratories do not have the capacity to diagnose, as PCR testing or special culturing is only available in research laboratories and for those laboratories that can then they have severely restricted the screening of respiratory viruses, unless specifically requested.

This is mainly due to long standing concerns about self limiting infections, costs, biosafety level requirements, competing health priorities and turnaround time for reporting, which delays meaningful diagnosis that needs to be made early enough to affect patient management. However, considering the number of hospital visits and lost school days due to respiratory tract infections this information should be utilized to formulate appropriate informative differential laboratory diagnostic approaches to respiratory infections especially in children and should mount to provision of better management of these infections.

Thirdly, there is more than one species of HBoV. We have only identified one. More epidemiological studies employing full genomic sequencing strategies should be initiated for this virus and other novel emerging viruses of the human respiratory tract.

Finally, as envisioned by van den Hoogen (2001) it is evident that additional undiscovered respiratory pathogens are likely to exist. In view of current technological advancement, viral studies are likely to experience rapid growth driven by improved and newer molecular assay techniques, reduced cost of DNA sequencing and more importantly advances in bioinformatics. We envision the exponential isolation and identification of new viral agents worldwide at a rate considerably greater than the knowledge of their biology.

6.0 References

Allander, T. (2008). Human bocavirus. J Clin Virol **41**(1): 29-33.

Allander, T., K. Andreasson, et al. (2007a). Identification of a third human polyomavirus. J Virol **81**(8): 4130-4136.

Allander, T., M. T. Tammi, et al. (2005). Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proc Natl Acad Sci U S A **102**(36): 12891-6.

Allander, T., T. Jartti, et al. (2007b). Human bocavirus and acute wheezing in children. Clin Infect Dis **44**(7): 904-10.

Arden, K. E., P. McErlean, et al. (2006). Frequent detection of human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute respiratory tract infections. J Med Virol **78**(9): 1232-40.

Bastien, N., K. Brandt, et al. (2006). Human Bocavirus infection, Canada. Emerg Infect Dis **12**(5): 848-50.

Beder, L. B., M. Hotomi, et al. (2009). Clinical and microbiological impact of human bocavirus on children with acute otitis media. Eur J Pediatr.

Benson, D. A., I. Karsch-Mizrachi, et al. (2010). GenBank." Nucleic Acids Res **38**(Database issue): D46-51.

Blessing, K., F. Neske, et al. (2009). Prolonged detection of human bocavirus DNA in nasopharyngeal aspirates of children with respiratory tract disease. Pediatr Infect Dis J **28**(11): 1018-1019.

Brieu, N., B. Gay, et al. (2007). Electron microscopy observation of human bocavirus (HBoV) in nasopharyngeal samples from HBoV-infected children. J Clin Microbiol **45**(10): 3419-20.

Chieochansin, T., S. Chutinimitkul, et al. (2007). Complete coding sequences and phylogenetic analysis of Human Bocavirus (HBoV). Virus Res **129**(1): 54-7.

Choi, E. H., H. J. Lee, et al. (2006). The association of newly identified respiratory viruses with lower respiratory tract infections in Korean children, 2000-2005. Clin Infect Dis **43**(5): 585-92.

Chung, J. Y., T. H. Han, et al. (2008). Th1 and Th2 cytokine levels in nasopharyngeal aspirates from children with human bocavirus bronchiolitis. J Clin Virol **43**(2): 223-5.

Clement, N., G. Battaglioli, et al. (2009). Prevalence of human bocavirus in human tonsils and adenoids. Emerg Infect Dis **15**(7): 1149-50.

de Vries, J. J., R. G. Bredius, et al. (2009). Human bocavirus in an immunocompromised child presenting with severe diarrhoea. J Clin Microbiol **47**(4): 1241-3.

Dijkman, R., S. M. Koekkoek, et al. (2009). Human bocavirus can be cultured in differentiated human airway epithelial cells. J Virol **83**(15): 7739-48.

Ditt, V., S. Viazov, et al. (2008). Genotyping of human bocavirus using a restriction length polymorphism. Virus Genes **36**(1): 67-9.

Esposito, S., S. Bosis, et al. (2008). Impact of human bocavirus on children and their families. J Clin Microbiol **46**(4): 1337-1342.

Feng, H., M. Shuda, et al. (2008). Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science **319**(5866): 1096-100.

Fischer, S. A. (2008). Emerging viruses in transplantation: there is more to infection after transplant than CMV and EBV. Transplantation **86**(10): 1327-39.

Foulongne, V., Y. Olejnik, et al. (2006). Human bocavirus in French children. Emerg Infect Dis **12**(8): 1251-3.

Fry, A. M., X. Lu, et al. (2007). Human bocavirus: a novel parvovirus epidemiologically associated with pneumonia requiring hospitalization in Thailand. J Infect Dis **195**(7): 1038-45.

Fryer, J. F., A. Kapoor, et al. (2006). Novel parvovirus and related variant in human plasma. Emerg Infect Dis **12**(1): 151-154.

Fryer, J. F., E. Delwart, et al. (2007a). Frequent detection of the parvoviruses, PARV4 and PARV5, in plasma from blood donors and symptomatic individuals. Transfusion **47**(6): 1054-61.

Fryer, J. F., S. B. Lucas, et al. (2007b). Parvoviruses PARV4/5 in hepatitis C virus-infected patient. Emerg Infect Dis **13**(1): 175-176.

Gregory, T. (2008). Understanding Evolutionary Trees. *Evolution: Education and Outreach, 1*(2), 121-137.

Hamza, I. A., L. Jurzik, et al. (2009). Detection and Quantification of Human Bocavirus in River Water. J Gen Virol.

Hazlett, D. T., T. M. Bell, et al. (1988). Viral etiology and epidemiology of acute respiratory infections in children in Nairobi, Kenya. Am J Trop Med Hyg **39**(6): 632-640.

Hindiyeh, M. Y., N. Keller, et al. (2008). High rate of human bocavirus and adenovirus coinfection in hospitalized Israeli children. J Clin Microbiol **46**(1): 334-7.

Kainulainen, L., M. Waris, et al. (2008). Hepatitis and human bocavirus primary infection in a child with T-cell deficiency. J Clin Microbiol **46**(12): 4104-5.

Kantola, K., L. Hedman, et al. (2008). Serodiagnosis of human bocavirus infection. Clin Infect Dis **46**(4): 540-6.

Kaplan, N. M., W. Dove, et al. (2006). Human bocavirus infection among children, Jordan. Emerg Infect Dis **12**(9): 1418-20.

Kapoor, A., E. Slikas, et al. (2009). A newly identified bocavirus species in human stool. J Infect Dis **199**(2): 196-200.

Kapoor, A., N. Mehta, et al. (2010a). Identification and characterization of a new bocavirus species in gorillas. PLoS One **5**(7): e11948.

Kapoor, A., P. Simmonds, et al. (2010b). "Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent in enteric infections." J Infect Dis **201**(11): 1633-1643.

Karalar, L., J. Lindner, et al. (2010). Prevalence and clinical aspects of human bocavirus infection in children. Clin Microbiol Infect **16**(6): 633-639.

Kenya Demographic Health Survey report 2008 - 2009

Kesebir, D., M. Vazquez, et al. (2006). Human bocavirus infection in young children in the United States: molecular epidemiological profile and clinical characteristics of a newly emerging respiratory virus. J Infect Dis **194**(9): 1276-82.

Kleines, M., S. Scheithauer, et al. (2007). High prevalence of human bocavirus detected in young children with severe acute lower respiratory tract disease by use of a standard PCR protocol and a novel real-time PCR protocol. J Clin Microbiol **45**(3): 1032-4.

Knipe DM, Peter M. Howley, et al (2007). Fields virology fifth edition. Lippincott Williams & Wilkins

Kupfer, B., J. Vehreschild, et al. (2006). Severe pneumonia and human bocavirus in adult. Emerg Infect Dis **12**(10): 1614-6.

Lau, S. K., C. C. Yip, et al. (2009). Clinical and molecular epidemiology of human rhinovirus C in children and adults in Hong Kong reveals a possible distinct human rhinovirus C subgroup. J Infect Dis **200**(7): 1096-103.

Lemey, P, Salemi M and Vandamme A (2009). The phylogenetic handbook. A practical approach to phylogenetic analysis and hypothesis testing. Cambridge university press, Newyork, USA

Lin, F., W. Guan, et al. (2008). ELISAs using human bocavirus VP2 virus-like particles for detection of antibodies against HBoV. J Virol Methods **149**(1): 110-7.

Lindner, J., L. Karalar, et al. (2008a). Clinical and epidemiological aspects of human bocavirus infection. J Clin Virol **43**(4): 391-5.

Lindner, J., L. Karalar, et al. (2008b). Humoral immune response against human bocavirus VP2 virus-like particles. Viral Immunol **21**(4): 443-9.

Longtin, J., M. Bastien, et al. (2008). Human bocavirus infections in hospitalized children and adults. Emerg Infect Dis **14**(2): 217-21.

Luna, L. K., M. Panning, et al. (2007). Spectrum of viruses and atypical bacteria in intercontinental air travellers with symptoms of acute respiratory infection. J Infect Dis **195**(5): 675-9.

Ma, X., R. Endo, et al. (2006). Detection of human bocavirus in Japanese children with lower respiratory tract infections. J Clin Microbiol **44**(3): 1132-4.

Maggi, F., E. Andreoli, et al. (2007). Human bocavirus in Italian patients with respiratory diseases. J Clin Virol **38**(4): 321-5.

Mahony, J. B. (2008). Detection of respiratory viruses by molecular methods. Clin Microbiol Rev **21**(4): 716-47.

Manning, A., V. Russell, et al. (2006). Epidemiological profile and clinical associations of human bocavirus and other human parvoviruses. J Infect Dis **194**(9): 1283-90.

Manteufel, J. and U. Truyen (2008). Animal bocaviruses: a brief review. Intervirology **51**(5): 328-34.

Marra, M. A., S. J. Jones, et al. (2003). The Genome sequence of the SARS-associated coronavirus. Science **300**(5624): 1399-1404.

McIntosh, K. (2006). Human bocavirus: developing evidence for pathogenicity. J Infect Dis **194**(9): 1197-9.

Monteny, M., H. G. Niesters, et al. (2007). Human bocavirus in febrile children, The Netherlands. Emerg Infect Dis **13**(1): 180-2.

Morse, S. S. (1995). Factors in the emergence of infectious diseases. Emerg Infect Dis **1**(1): 7-15.

Naghipour, M., L. E. Cuevas, et al. (2007). Human bocavirus in Iranian children with acute respiratory infections. J Med Virol 79(5): 539-43.

Neske, F., K. Blessing, et al. (2007). Real-time PCR for diagnosis of human bocavirus infections and phylogenetic analysis. J Clin Microbiol 45(7): 2116-22.

Nishizawa, T., H. Okamoto, et al. (1997). A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown aetiology. Biochem Biophys Res Commun 241(1): 92-7.

Pozo, F., M. L. Garcia-Garcia, et al. (2007). High incidence of human bocavirus infection in children in Spain. J Clin Virol 40(3): 224-8.

Qu, X. W., W. P. Liu, et al. (2008). Phospholipase A2-like activity of human bocavirus VP1 unique region. Biochem Biophys Res Commun 365(1): 158-63.

Qu, X. W., Z. J. Duan, et al. (2007). Human bocavirus infection, People's Republic of China. Emerg Infect Dis 13(1): 165-8.

Regamey, N., U. Frey, et al. (2007). Isolation of human bocavirus from Swiss infants with respiratory infections. Pediatr Infect Dis J 26(2): 177-9.

Schildgen, O., A. Muller, et al. (2008). Human bocavirus: passenger or pathogen in acute respiratory tract infections? Clin Microbiol Rev 21(2): 291-304, table of contents.

Schwartz, D., B. Green, et al. (2002). The canine minute virus (minute virus of canines) is a distinct parvovirus that is most similar to bovine parvovirus. Virology 302(2): 219-23.

Shackelton, L. A., K. Hoelzer, et al. (2007). Comparative analysis reveals frequent recombination in the parvoviruses. J Gen Virol 88(Pt 12): 3294-3301.

Simon, A., P. Groneck, et al. (2007). Detection of bocavirus DNA in nasopharyngeal aspirates of a child with bronchiolitis. J Infect 54(3): e125-7.

Simons, J. N., T. P. Leary, et al. (1995). Isolation of novel virus-like sequences associated with human hepatitis. Nat Med 1(6): 564-9.

Sloots, T. P., D. M. Whiley, et al. (2008). Emerging respiratory agents: new viruses for old diseases? J Clin Virol 42(3): 233-43.

Sloots, T. P., P. McErlean, et al. (2006). Evidence of human coronavirus HKU1 and human bocavirus in Australian children. J Clin Virol 35(1): 99-102.

Smuts, H. and D. Hardie (2006). Human bocavirus in hospitalized children, South Africa. Emerg Infect Dis 12(9): 1457-8.

Smuts, H., L. Workman, et al. (2008). Role of human metapneumovirus, human coronavirus NL63 and human bocavirus in infants and young children with acute wheezing. J Med Virol 80(5): 906-12.

Soderlund-Venermo, M., A. Lahtinen, et al. (2009). Clinical assessment and improved diagnosis of bocavirus-induced wheezing in children, Finland. Emerg Infect Dis 15(9): 1423-30.

Tamura, K., J. Dudley, et al. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24(8): 1596-9.

Urisman, A., R. J. Molinaro, et al. (2006). Identification of a novel Gammaretrovirus in prostate tumours of patients homozygous for R462Q RNASEL variant. PLoS Pathog 2(3): e25.

Vallet, C., C. Pons-Catalano, et al. (2009). Human bocavirus: a cause of severe asthma exacerbation in children. J Pediatr 155(2): 286-8.

Van den Hoogen, B. G., J. C. de Jong, et al. (2001). A newly discovered human pneumovirus isolated from young children with respiratory tract disease. Nat Med 7(6): 719-24.

Van der Hoek, L., K. Pyrc, et al. (2004). Identification of a new human coronavirus. Nat Med 10(4): 368-73.

Van Regenmortel, M. H. (2007). Virus species and virus identification: past and current controversies. Infect Genet Evol 7(1): 133-144.

Villa, L., S. Melon, et al. (2008). Detection of human bocavirus in Asturias, Northern Spain. Eur J Clin Microbiol Infect Dis 27(3): 237-9.

Weissbrich, B., F. Neske, et al. (2006). Frequent detection of bocavirus DNA in German children with respiratory tract infections. BMC Infect Dis 6: 109.

Williams, J. V. (2010). Deja vu all over again: Koch's postulates and virology in the 21st century. J Infect Dis 201(11): 1611-1614.

Woo, P. C., S. K. Lau, et al. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia." J Virol 79(2): 884-895.

Zehender, G., C. De Maddalena, et al. (2010). Rapid molecular evolution of human bocavirus revealed by Bayesian coalescent inference. Infect Genet Evol 10(2): 215-20.

7.0 Appendix

1a. Isolate KE-10N8/3, 762bp, plus strand

```
1   CAATGAAAGA GCATACATTC CTCCTGGACT AATGTTTAAT CCAAAAGTTC CAACAAGAAG
61  AGTTCAGTAC ATAAGACAAA ACGGAAGCAC AGCAGCCAGC ACAGGCAGAA TTCAGCCATA
121 CTCAAAACCA ACAAGCTGGA TGACAGGACC TGGCCTGCTC AGTGCACAGA GAGTAGGACC
181 ACAGTCATCA GACACTGCTC CATTCATGGT TTGCACTAAC CCAGAAGGAA CACACATAAA
241 CACAGGTGCT GCAGGATTTG GATCTGGCTT TGATCCTCCA AGCGGATGTC TGGCACCAAC
301 TAACCTAGAA TACAAACTTC AGTGGTACCA GACACCAGAA GGAACAGGAA ATAATGGAAA
361 CATAATTGCA AACCCATCAC TCTCAATGCT TAGAGACCAA CTCCTATACA AAGGAAACCA
421 GACCACATAC AATCTAGTGG GGGACATATG GATGTTTCCA AATCAAGTCT GGGACAGATT
481 TCCTATCACC AGAGAAAATC CAATCTGGTG CAAAAAACCA AGGGCTGACA AACACACAAT
541 CATGGATCCA TTTGATGGAT CAATTGCAAT GGATCATCCT CCAGGCACTA TTTTTATAAA
601 AATGGCAAAA ATTCCAGTTC CAACTGCCTC AAATGCAGAC TCATACCTAA ACATATACTG
661 TACTGGACAA GTCAGCTGTG AGATTGTATG GGAAGTAGAA AGATACGCAA CAAAGAACTG
721 GCGTCCAGAA AGAAGACATA CTGCACTCGG GATGTCACTG GG
```

1b. Isolate KE- 9N7/, 776bp, plus strand

```
1     GAAATTTACT TTTAACTTTG ACTGTGAATG GGGTTAACAA TGAAAGAGCA TACATTCCTC
61  CTGGACTAAT GTTTAATCCA AAAGTTCCAA CAAGAAGAGT TCAGTACATA AGACAAAACG
121 GAAGCACAGC AGCCAGCACA GGCAGAATTC AGCCATACTC AAAACCAACA AGCTGGATGA
181 CAGGACCTGG CCTGCTCAGT GCACAGAGAG TAGGACCACA GTCATCAGAC ACTGCTCCAT
241 TCATGGTTTG CACTAACCCA GAAGGAACAC ACATAAACAC AGGTGCTGCA GGATTTGGAT
301 CTGGCTTTGA TCCTCCAAGC GGATGTCTGG CACCAACTAA CCTAGAATAC AAACTTCAGT
361 GGTACCAGAC ACCAGAAGGA ACAGGAAATA ATGGAAACAT AATTGCAAAC CCATCACTCT
421 CAATGCTTAG AGACCAACTC CTATACAAAG GAAACCAGAC CACATACAAT CTAGTGGGGG
481 ACATATGGAT GTTTCCAAAT CAAGTCTGGG ACAGATTTCC TATCACCAGA GAAAATCCAA
541 TCTGGTGCAA AAAACCAAGG GCTGACAAAC ACACAATCAT GGATCCATTT GATGGATCAA
601 TTGCAATGGA TCATCCTCCA GGCACTATTT TTATAAAAAT GGCAAAAATT CCAGTTCCAA
661 CTGCCTCAAA TGCAGACTCA TACCTAAACA TATACTGTAC TGGACAAGTC AGCTGTGAGA
721 TTGTATGGGA AGTAGAAAGA TACGCAACAA AGAACTGGCG TCCAGAAAGA AGACAT
```

1c. Isolate KE-23A8/11, 790bp, plus strand

```
1    CACTGAAATT TACTTTTAAC TTTGACTGTG AATGGGTTAA CAATGAAAGA GCATACATTC
61   CTCCTGGACT AATGTTTAAT CCAAAAGTTC CAACAAGAAG AGTTCAGTAC ATAAGACAAA
121 ACGGAAGCAC AGCAGCCAGC ACAGGCAGAA TTCAGCCATA CTCAAAACCA ACAAGCTGGA
181 TGACAGGACC TGGCCTGCTC AGTGCACAGA GAGTAGGACC ACAGTCATCA GACACTGCTC
241 CATTCATGGT TTGCACTAAC CCAGAAGGAA CACACATAAA CACAGGTGCT GCAGGATTTG
301 GATCTGGCTT TGATCCTCCA AGCGGATGTC TGGCACCAAC TAACCTAGAA TACAAACTTC
361 AGTGGTACCA GACACCAGAA GGAACAGGAA ATAATGGAAA CATAATTGCA AACCCATCAC
421 TCTCAATGCT TAGAGACCAA CTCCTATACA AAGGAAACCA GACCACATAC AATCTAGTGG
481 GGGACATATG GATGTTTCCA AATCAAGTCT GGGACAGATT TCCTATCACC AGAGAAAATC
541 CAATCTGGTG CAAAAAACCA AGGGCTGACA AACACACAAT CATGGATCCA TTTGATGGAT
601 CAATTGCAAT GGATCATCCT CCAGGCACTA TTTTTATAAA AATGGCAAAA ATTCCAGTTC
661 CAACTGCCTC AAATGCAGAC TCATACCTAA ACATATACTG TACTGGACAA GTCAGCTGTG
721   AGATTGTATG   GGAAGTAGAA   AGATACGCAA   CAAAGAACTG   GCGTCCAGAA
AGAAGACATA
781 CTGCACTCGG
```

1d. Isolate KE- 48G7/8, 733bp, plus strand

```
1    CACTGAATTT ACTTTTTAAC TTTGACTGTG AATGGGTTAA CAATGAAAAG AGCATACATT
61   CCTCCTGGAC TAATGTTTAA TCCAAAAGTT CCAACAAGAA GAGTTCAGTA CATAAGACAA
121 AACGGAAGCA CAGCAGCCAG CACAGGCAGA ATTCAGCCAT ACTCAAAACC AACAAGCTGG
181 ATGACAGGAC CTGGCCTGCT CAGTGCACAG AGAGTAGGAC CACAGTCATC AGACACTGCT
241CCATTCATGG TTTGCACTAA CCCAGAAGGA ACACACATAA ACACAGGTGC TGCAGGATTT
301 GGATCTGGCT TTGATCCTCC AAGCGGATGT CTGGCACCAA CTAACCTAGA ATACAAACTT
361 CAGTGGTACC AGACACCAGA AGGAACAGGA AATAATGGAA ACATAATTGC AAACCCATCA
421 CTCTCAATGC TTAGAGACCA ACTCCTATAC AAAGGAAACC AGACCACATA CAATCTAGTG
481 GGGGACATAT GGATGTTTCC AAATCAAGTC TGGGACAGAT TTCCTATCAC CAGAGAAAAT
541 CCAATCTGGT GCAAAAAACC AAGGGCTGAC AAACACACAA TCATGGATCC ATTTGATGGA
601 TCAATTGCAA TGGATCATCC TCCAGGCACT ATTTTTATAA AAATGGCAAA AATTCCAGTT
661 CCAACTGCCT CAAATGCAGA CTCATACCTA AACATATACT GTACTGGGAC AAGTCAGCTG
721 TGAGATTGTA TGG
```

1e. Isolate KE-4S9/04, 749 bp DNA, plus strand

```
1   TTTAACTTTG ACTGTGAATG GGGTTAACAA TGAAAGAGCA TACATTCCTC CTGGACTAAT
61  GTTTAATCCA AAAAGTTCCA ACAAGAAGAG TTCAGTACAT AAGACAAAAC GGAAGCACAG
121 CAGCCAGCAC AGGCAGAATT CAGCCATACT CAAAACCAAC AAGCTGGATG ACAGGACCTG
181 GCCTGCTCAG TGCACAGAGA GTAGGACCAC AGTCATCAGA CACTGCTCCA TTCATGGTTT
241 GCACTAACCC AGAAGGAACA CACATAAACA CAGGTGCTGC AGGATTTGGA TCTGGCTTTG
301 ATCCTCCAAG CGGATGTCTG GCACCAACTA ACCTAGAATA CAAACTTCAG TGGTACCAGA
361 CACCAGAAGG AACAGGAAAT AATGGAAACA TAATTGCAAA CCCATCACTC TCAATGCTTA
421 GAGACCAACT CCTATACAAA GGAAACCAGA CCACATACAA TCTAGTGGGG GACATATGGA
481 TGTTTCCAAA TCAAGTCTGG GACAGATTTC CTATCACCAG AGAAAATCCA ATCTGGTGCA
541 AAAAACCAAG GGCTGACAAA CACACAATCA TGGATCCATT TGATGGATCA ATTGCAATGG
601 ATCATCCTCC AGGCACTATT TTTATAAAAA TGGCAAAAAT TCCAGTTCCA ACTGCCTCAA
661 ATGCAGACTC ATACCTAAAC ATATACTGTA CTGGACAAGT CAGCTGTGAG ATTGTATGGG
721 AAGTAGAAAG ATACGCAACA AAGAACTGG
```