DOCTORAL SCHOOL IN BIOLOGY

Biomedical Sciences and Technologies

CYCLE XXXI

Monitoring Hepatitis E Virus and Coronavirus in animal reservoirs by classical molecular techniques and Next Generation Sequencing

Luca De Sabato

Tutor
Dr. Ilaria Di Bartolo
Dr. Gabriele Vaccari
Prof. Elisabetta Affabris

Coordinator
Prof. Paolo Visca

Sommario

Il virus dell'Epatite E (HEV) e i Coronavirus sono virus a RNA a trasmissione zoonotica che negli ultimi anni hanno destato preoccupazione per la Sanità Pubblica poiché responsabili di patologie umane emergenti.

Il virus dell'Epatite E è un virus a trasmissione oro-fecale responsabile, nell'uomo, di forme di epatite acuta con scarsa tendenza alla cronicizzazione. Negli animali, al contrario, l'infezione evolve in maniera asintomatica. Nei paesi in via di sviluppo, il virus è endemico ed è responsabile di vaste epidemie conseguenti alla contaminazione fecale delle acque potabili. Nei paesi industrializzati, i casi umani autoctoni, non riconducibili quindi a soggiorni in aree endemiche, derivano dalla trasmissione zoonotica dei virus da serbatoi animali rappresentati prevalentemente dai suini (principale serbatoio), dai cinghiali e dal cervo. I genotipi 3 e 4 (HEV-3 e HEV-4) sono zoonotici e i ceppi umani e animali presentano fino al 100% d'identità nucleotidica. Nei paesi industrializzati HEV si trasmette prevalentemente attraverso il consumo di carne o prodotti a base di carne (in particolare salsicce contenenti fegato) consumati crudi o poco cotti. In Italia, il genotipo 3 è stato identificato sia in casi umani, sia nei suini allevati e nei cinghiali.

I Coronavirus infettano l'uomo e altre specie animali come pipistrelli, suini, bovini, gatti, ratti e molti altri ancora. Negli ultimi 10 anni due specie di Coronavirus sono risultate zoonotiche: il virus della Sindrome Respiratoria Acuta Grave (SARS) e il virus della Sindrome Respiratoria Medio Orientale (MERS) che hanno causato migliaia di casi di infezione e di morti in tutto il mondo. Studi filogenetici sui Coronavirus animali e umani hanno dimostrato che molte delle specie circolanti hanno avuto origine da ceppi circolanti nei pipistrelli, considerati il serbatoio principale. Il ciclo epidemiologico ha un ospite intermedio, la civetta delle palme per il SARS-CoV e i dromedari per il MERS-CoV, che ha il solo ruolo di amplificare la popolazione virale prima dell'adattamento finale del virus all'uomo. Dopo le epidemie di SARS e MERS, la sorveglianza dei coronavirus nei pipistrelli è aumentata in tutto il mondo. I coronavirus sono stati identificati anche nei pipistrelli Europei e Italiani rilevando la presenza di ceppi correlati sia a specie di coronavirus non zoonotici, sia a virus simili al SARS-CoV e MERS-CoV confermando i pipistrelli come animali serbatoio per questi ceppi.

Il virus dell'Epatite E e alcune specie di Coronavirus sono ancora poco caratterizzati, in parte perché identificati di recente e in parte perché manca un sistema efficiente di crescita *in vitro*.

Le informazioni genetiche disponibili su questi virus sono il risultato delle analisi filogenetiche condotte sulle sequenze virali necessarie a stabilire le relazioni evolutive tra i virus e la loro classificazione in specie per i Coronavirus e in genotipi ulteriormente divisi in sottotipi per HEV. Tuttavia, le sequenze disponibili sono ancora poche e per lo più rappresentate da porzioni corte del genoma e talvolta poco informative. Il nuovo metodo di sequenziamento chiamato Sequenziamento di Nuova Generazione – (NGS) offre la possibilità di acquisire un numero elevato di sequenze tali da

permettere di ricostruire i genomi interi, con notevole risparmio di tempo e in modo indipendente dalla conoscenza a priori della specie virale coinvolta.

Il principale obiettivo di questo lavoro è stato la ricerca dei virus dell'Epatite E e di Coronavirus in alcuni serbatoi animali, la caratterizzazione dei ceppi circolanti mediante sequenziamento e lo studio delle loro relazioni evolutive. Il protocollo di NGS utilizzato è stato sviluppato durante lo studio, sia nella parte di preparazione dei campioni sia nello sviluppo delle pipeline bioinformatiche per le analisi dei risultati.

Il protocollo NGS applicato a due campioni di feci di suino positive per HEV-3 non è stato sufficientemente sensibile per poter ottenere un genoma completo su una matrice complessa come le feci. La sequenza del genoma è stata completata mediante tecniche convenzionali. I genomi completi dei due ceppi di HEV-3 sono stati caratterizzati mediante filogenesi, permettendo di stabilire la loro appartenenza a un nuovo sottotipo di HEV-3 denominato HEV-3l. I ceppi appartenenti a questo nuovo sottotipo sono stati identificati a oggi unicamente in Italia e in Francia e potrebbero quindi rappresentare un sottotipo virale che non si è adattato al suino in maniera definitiva e che ancora circola moderatamente nella popolazione suina.

Un approccio simile è stato applicato per lo studio di ceppi di HEV-3 identificati in fegati di cinghiale. Nella fase preliminare di questo lavoro è stato condotto uno studio di sorveglianza sulla circolazione di HEV in popolazioni di cinghiali cacciati in alcune aree del sud dell'Italia. I ceppi di HEV-3 identificati (prevalenza media 13.7%) sono risultati, da una prima analisi di sequenza su regioni genomiche corte, non classificabili nei sottotipi di HEV-3 noti fino a quel momento. Gli RNA virali estratti dai fegati di cinghiale positivi per HEV-3 sono stati sequenziati mediante NGS. I genomi completi (7,2 kb) ottenuti sono stati analizzati mediante filogenesi. Uno dei ceppi ha mostrato un basso grado di relazione con altri ceppi di riferimento, appartenendo probabilmente a un nuovo sottotipo. Il secondo ceppo è stato invece classificato in un sottotipo già rilevato in precedenza in Europa nei cinghiali e in casi umani ma mai rilevato in Italia.

Un secondo studio di sorveglianza molecolare è stato condotto su alcune popolazioni di cinghiali cacciati nella provincia di Viterbo. Il 52,2% dei fegati esaminati è risultato positivo alla ricerca del genoma di HEV-3. Lo studio di sequenza condotto sui ceppi identificati negli animali cacciati in diverse zone della provincia ha evidenziato la circolazione di più ceppi virali geneticamente distanti tra loro. Tuttavia, un unico ceppo virale (100% di identità nucleotidica) circolava tra gli animali appartenenti allo stesso gruppo familiare, cacciati dalla stessa squadra nel corso di uno specifico giorno di attività venatoria. In conclusione, i risultati ottenuti evidenziano il potenziale rischio di trasmissione zoonotica confermato dalla stretta vicinanza genetica tra i ceppi virali identificati nei serbatoi animali e quelli che sono stati identificati in casi umani. Per caratterizzare le specie di Coronavirus circolanti nelle popolazioni di pipistrelli in Italia il primo protocollo di NGS è stato

applicato per il sequenziamento di ceppi identificati in pipistrelli provenienti dall'Emilia-Romagna. Il protocollo NGS è stato applicato all'RNA estratto da 3 campioni fecali e 2 carcasse positivi, rispettivamente, per Alpha e Betacoronavirus, ottenendo la sequenza completa dei genomi virali (30kb). I due ceppi virali identificati nella specie *Pipistrellus kuhlii* e *Hypsugo savii* sono stati classificati in un'unica specie e hanno mostrato relazioni evolutive con i ceppi di MERS-CoV confermando la circolazione anche in specie di pipistrelli presenti nel territorio italiano. Inoltre la presenza di un'unica specie (MERS-CoV) di Coronavirus in due pipistrelli di specie e genere diversi suggerisce l'adattamento del virus a due specie animali che occupano situazioni ambientali differenti essendo i pipistrelli *Pipistrellus kuhlii* legati alle aree agricole ed urbane mentre le specie *Hypsugo savii* a quelle rurali. I tre ceppi rilevati nei pipistrelli della specie *Pipistrellus kuhlii* sono stati invece classificati in due nuove specie di coronavirus mai rilevate in precedenza e hanno mostrato elevata divergenza nucleotidica con ceppi di coronavirus umani e di pipistrello. Queste analisi hanno confermato l'eterogeneità dei coronavirus confermando la presenza di due specie virali differenti in grado di infettare la stessa specie di pipistrelli.

Per migliorare il protocollo di NGS cercando di ottenere una sensibilità maggiore e una maggiore accuratezza del risultato attraverso un aumento della profondità di sequenziamento, è stato sviluppato un secondo metodo NGS basato sull'amplificazione dell'RNA virale target mediante coppie di primer disegnate sul genoma di diversi ceppi e sottotipi di HEV-3 e sulle sequenze della specie MERS-CoV. Il nuovo metodo ha permesso di ottenere il genoma completo di cinque ceppi virali appartenenti a diversi sottotipi di HEV-3, distanti tra di loro e scelti dal precedente studio di sorveglianza di HEV sui fegati di cinghiale. L'analisi filogenetica condotta sui genomi completi ha permesso di classificare due ceppi in sottotipi HEV-3 già noti e comuni nelle popolazioni di cinghiali. Mentre gli altri due ceppi hanno mostrato una scarsa correlazione con i ceppi HEV-3 noti e potrebbero rappresentare nuovi sottotipi di HEV-3.

Il nuovo metodo è stato testato inizialmente su un isolato cellulare di MERS-CoV permettendo il sequenziamento del genoma virale completo del ceppo. Per valutare la robustezza del metodo, questo è stato applicato su campioni di campo provenienti da pipistrelli. In questo reservoir sono stati ad oggi identificati solo ceppi correlati, definiti MERS-CoV like, ma non identici al MERS-CoV identificato nell'uomo. Il protocollo ha permesso il sequenziamento del genoma completo dei ceppi coinvolti nonostante la divergenza genetica con quello del MERS-CoV.

Questo nuovo approccio ha dato risultati ottimali, infatti sono stati ottenuti i genomi completi sia di ceppi di HEV-3 appartenenti a diversi sottotipi sia di ceppi appartenenti alla specie MERS-CoV tra loro divergenti. Questo fa pensare al possibile utilizzo del pannello di primer e del metodo anche per la diagnostica di questi virus. Inoltre, rispetto al protocollo applicato in precedenza, l'accuratezza del risultato è stata elevata consentendo di ottenere un numero elevato di sequenze specifiche virali.

Il lavoro descritto in questa tesi ha avuto l'obiettivo generale di monitorare e caratterizzare ceppi di virus zoonotici (HEV) o potenzialmente zoonotici (Coronavirus) circolanti, in Italia, in alcune specie animali e valutarne le relazioni evolutive mediante analisi filogenetiche sulle sequenze. I risultati hanno permesso di identificare nuovi sottotipi del genotipo zoonotico HEV-3 e nuove specie di Coronavirus mai identificati in precedenza. Inoltre, il metodo di NGS sviluppato (sequenziamento e pipeline bionformatica) ha dato risultati ottimali, migliorabili nella fase iniziale di arricchimento delle sequenze virali e sarà utile nella sorveglianza in ambito animale e umano delle infezioni da parte di HEV e MERS-CoV.

Table of contents

CHAPTER 1

Introduction

1. Introduction

The World Health Organization (WHO) defined the zoonosis as an infection disease transmitted from animals to humans by bacterial, viral, parasites or unconventional pathogens (http://www.who.int). Up to 75% of "emerging" diseases affecting humans are zoonoses (Taylor et al., 2001) cause by pathogenic agents infecting naturally wild or domestic non-human reservoirs asymptomatically. The climate changes, habitat modifications (intensive farming, modern agriculture, humans encroachment), demographic changes increase the interactions between humans and wild or domestic animals and the emerging or re-emerging of zoonotic diseases. Among zoonotic agents, viruses constitute about 60% of the known human pathogens, and RNA-viruses are the 35% of them (Taylor et al., 2001; Woolhouse & Gowtage-Sequeria, 2005; Kilpatrick & Randolph, 2012). Due to the lack of proofreading activity of RNA virus polymerases, the event of recombination and reassortment of genomes, new viral genetic variants are constantly generated (Manrubia, 2006). Consequently RNA viruses evolve fast and the existence of multiple variant strains provides to RNA viruses the capacity to adapt to host and environmental changes, increasing transmission of the pathogen and spill-over into humans. Quick and accurate diagnostic techniques are required to prevent and respond to emerging or re-emerging zoonotic pathogens. The diagnostic of viral infection is now mainly based on polymerase chain reaction (PCR), its evolution (Real-Time PCR) and the genomic sequencing, allowing the direct characterization of viral agents and overcoming the limits of some classical approaches. However, novel viruses cannot be either detected or characterized by traditional technique and molecular biology (Bexfield & Kellam, 2011). In 2000, the Next Generation Sequencing (NGS) technology was introduced for the first time (Brenner et al., 2000). This technique produces a great amount of sequences in a single experiment, over a relative short time and with reasonable cost. This approach has been applied in different fields such as clinical microbiology, oncology, nutrition, and pharmaceutics (Barzon et al., 2011). In microbiology, NGS not only provides great amounts of sequences for genetic profiling, acquisition of entire genomes (shotgun sequence), investigation of the biodiversity in biological and environmental samples (metagenomics) and the detection of known pathogen variants (quasi-species, mutants with increased virulence), but it also allows the discovering of novel infectious agents. The protocols for sequencing and analyses of the results on virus are still poorly efficient, and the use of NGS for sequencing of virus full genomes is still difficult. Limited amount of virus in biological samples and problems with cultivating viruses largely hampered the final results. Furthermore, the bioinformatics tools and pipelines are still keeping on developing for novel or not well known viruses yet. The aim of the thesis was the investigation of the occurrence of zoonotic RNA viruses (Hepatitis E virus and coronavirus) in animal

hosts, studying their viral diversity and evolutive correlation by using conventional approaches supported by NGS sequencing.

1.1 Hepatitis E

Hepatitis E is an acute liver disease causing around 20 million infections worldwide and leading to an estimated 3.3 million symptomatic cases. The disease causes large outbreaks in low income countries, associated to consumption of drinking water contaminated by human feces. In developed countries, the disease causes sporadic cases or small outbreaks and the main route of transmission is foodborne. In Europe, in the last 10 years an unexplained number of hepatitis E case, the occurrence of chronic disease in transplanted patients and the wide presence of HEV in pigs and wild boar have been observed and led to a major concern on the disease that is now considered emerging.

1.1.1 Hepatitis E virus

The infection is caused by a small RNA virus named Hepatitis E virus (HEV). It has primarily been associated to large human outbreaks occurred in India and China since 1955 (Kamar et al., 2012). In 1983, the HEV viral particles were detected by electron microscopy (EM) in feces collected by a volunteer who had been orally administrated with fecal material from a patient with a suspected non-A, non-B, non-C hepatitis (Balayan et al., 1983), confirming that HEV was the etiological agent of the disease that was transmitted by oral-fecal route. HEV is a quasi-enveloped virus of 32-34 nm, composed by the assembling of 60 copies of the capsid protein (pORF2, 72 kDa) in a multimeric structure. Truncated ORF2 gene expression in baculovirus vectors generates virus-like particles (VLP) which is strongly immunogenic and resembles the native virus in the three-dimensional structure (Lin et al., 2016).

The virus, observed by EM, is shed into feces as naked virions. HEV in the blood is associated to a lipid membrane (Nagashima et al., 2017). The virion contains a single-stranded positive sense RNA of approximately 7.2 kilobases, organized in 5' terminus untranslated region, three Open Reading Frames (ORFs) partially overlapping, and the 3' terminus untranslated region.

The ORF1 (110 kDa) encodes the non-structural polyprotein of about 1691-1708 amino acids involved in virus replication and post-translational maturation of viral proteins. The functional domains are: methyltransferase (MetT), Y domain, papain-like cysteine protease (PCP), a proline rich region containing the hypervariable region (HVR), a macro domain (X domain), RNA helices and RNA-dependent RNA polymerase (RdRp). The ORF2 encodes the capsid protein of approximately 660 amino acids divided into three functional domains: S domain, M domain, and P domain. The ORF3 (13 kDa) encodes a phosphoprotein of approximately 360 amino acids, which function is not totally clear but is probably involved in the virion release from the cells and in the modulation of the immune responses in the acute phase (Cao & Meng, 2012).

HEV replication cycle

The mechanism of the life cycle of HEV is largely unknown due to the lack of a robust cell-culture system for virus replication. HEV virions exist in two forms: the naked virion (non-enveloped HEV) and the quasi-enveloped form (eHEV). Despite eHEV has no viral proteins on the surface, it remains infectious and is likely responsible for cell-to-cell spread within the host. The two forms probably enter in target cells by different mechanisms, but it is not clear yet. The non-enveloped HEV cell attachment is mediated by heparan sulfate proteoglycans (HSP). The eHEV attachment to the cells does not require HSP (Yin et al., 2016).

Some *in vitro* experiments, conducted with recombinant capsid protein (VLP) and cell culture assays, proved the binding of the C-terminal region of ORF2 to heat shock cognate protein 70 (HSC70) and the heparin sulfate proteoglycans HSPGs on the cell surface. After the binding to cells via heparan sulfate proteoglycans (HSPGs), HEV-VLPs are internalized via a dynamin-2-, clathrin-, and membrane cholesterol-dependent pathway (Yin et al., 2016).

The M domain in the ORF2 is conserved between all strains infecting mammalian and may be the putative binding motif. The proposed model for HEV replication is based on similarities with other well-characterized positive single-stranded viruses. Once penetrated into the permissive cells, the capsid protein may interact with heat shock protein 90 (HSP90) and glucose-regulated protein 78 (Grp78) for the intracellular transport and the HEV genomic RNA is uncoated (Cao & Meng, 2012). The 40S ribosomal subunit recognizes the 7-methylguanosine cap structure in the 5' UTR to initiate cap-dependent translation of viral proteins and the ORF1 is immediately translated into the cytoplasm of the infected cells to produce the non-structural polyproteins. The polyprotein undergoes cutting by the cellular proteases, also helped by the presence of the viral PCP. RdRp binds the cis-reactive elements (CRE) in the 3' UTR of the HEV genome and replicates the positive genomic filament in an intermediate negative strand for the production of positive-sense. The structural capsid protein and the ORF3 protein are translated from the bicistronic subgenomic RNA (Cao & Meng, 2012). The ORF2 contains a signal sequence at N- terminal for the translocation to the endoplasmatic reticulum (ER) where is glycosylated. In the HEV replication, the protein ORF3 seems to be involved in viral egress. The ORF3 interacts with the tumor susceptibility gene 101 (Tsg101) protein, a component of the cellular endosomal sorting complexes required for transport (ESCRT) machinery, promoting the budding of newly assembled HEV virions into multivesicular bodies (MVB) . In addition, ORF3 protein and markers for exosomes such as CD63 and CD81 are only found in eHEV. The release mechanism and the origin of envelope membrane remain uncertain. The main organ of HEV replication is liver, but various sites of extrahepatic replication have also been identified, including the intestine, lymph nodes, colon, spleen, peripheral blood monocytes in experimental infections using animal models (Yin et al., 2016).

HEV taxonomy and classification

HEV is the sole member of the *Hepeviridae* family. Two genera are included: the *Orthohepevirus* strains infect mammalian and avian species while the *Piscihepevirus* strains infect trout only (ICTV, https://talk.ictvonline.org). Besides the high genetic variability observed among HEV strains, a single serotype is recognized. *Orthohepevirus* includes four species (*A*, *B*, *C*, *D*) divided into different genotypes (Kamar et al., 2017). *Orthohepevirus A* genus includes strains grouped in 7 genotypes: HEV-1 and HEV-2 infect only humans, HEV-3 and HEV-4 are zoonotic and infect humans and suidae (pigs, wild boar) and circulate worldwide. HEV-5 and HEV-6 have only been detected in wild boars in Japan. HEV-7 has been identified in dromedary camels and in one patient affected by chronic hepatitis after liver transplantation. The HEV-8 has been recently identified in Bactrian camels (Lee et al., 2016; Woo et al., 2016). The *Orthohepevirus B* only infects avian species. *Orthohepevirus C* infects rat, greater bandicoot rat, Asian musk shrew, ferret and mink. *Orthohepevirus D* only infects bats (Smith et al., 2014). HEV strains identified in moose and kestrel have not been classified yet (Doceul et al., 2016) (Fig. 1).

Criteria established by ICTV for HEV classification are based on phylogenetic analysis and amino acid sequence identities comparisons of three HEV subgenomic regions: methyltransferase (ORF1-28 to 389 aa), helicase (ORF1-971 to 1185 aa) and replicase (ORF1-1249 to 1671 aa). Piscihepeviruses share with *Orthohepevirus* strains less than 30% amino acid identities in ORF1, ORF2 and ORF3 regions. In addition, they have a different genome organization with ORF3 displaced within the ORF2 (Smith et al., 2014). Host variability observed on the different *Orthohepevirus* genus is also confirmed by a great genetic variability. Phylogenetic analysis on HEV sequences confirmed the clustering of *Orthohepevirus* strain sequences into four distinct clades: *Orthohepevirus A, B, C* and *D*. They share amino acid sequence similarities of 42-49% (ORF1), 42-55% (ORF2) and 20-29% (ORF3) in the three HEV genomic regions. The 7 *Orthohepevirus A* genotypes have been defined by evaluating the amino acid distances of concatenated ORF1 and ORF2 (lacking hypervariable region) and using 0.088 p-distance value as a threshold to demarcate intra- and inter- genotype distances (Smith et al., 2014).

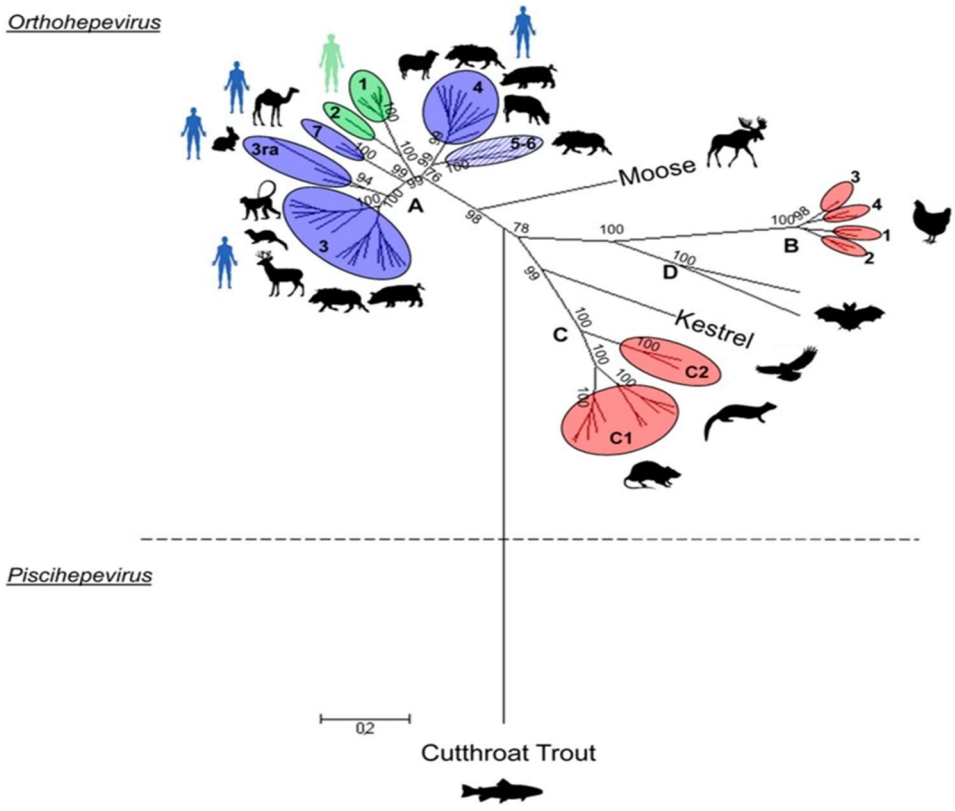
The strains infecting humans were initially classified into 4 genotypes and 24 subtypes based on the analysis of subgenomic regions (HEV-1, -2, -3, -4) (Lu et al., 2006). Recently, the HEV classification and the subtypes division have been updated by the analysis of the full genomes of HEV strains establishing 7 genotypes and 32 subtypes. To classify HEV strains in a defined subtype a list of reference complete genome sequences has been released (Smith et al., 2016). The classification or assignment into a specific subtype is based on the full genome sequences comparison with a set of HEV reference sequences and the phylogenetic analysis by clustering within subtype reference strains. Novel and unrelated strains, phylogenetically distinct from reference sequences, may be

assigned to a new subtype if at least three epidemiologically unrelated (originated from different studies, host or localities) complete genome sequences are available. In addition, the nucleotide distance values support the subtyping classification. However, the threshold value to discriminate between and within subtypes has not been established yet (Smith et al., 2016). Over the last 5 years, due to increased surveillance of the virus and the improving of the diagnostic techniques, the number of HEV sequences is continuously increasing and novel strains are being proposed, but a robust classification is still missing because of limited number of full genome sequences.

Orthohepevirus A genotypes are subdivided into subtypes based on the clustering observed by phylogenetic analysis on the alignment of complete genomes and the nucleotide p-distance between strains. The HEV-1 genotype is classified into 6 subtypes (a-f) divided into two clades: abcf and de. HEV-2 is represented by two subtypes 2a and 2b. The HEV-3 strains are classified into 11 subtypes (named a-k) and some unclassified strains and are grouped into 2 clades: abchjki and efg (Smith et al., 2014; Woo et al., 2016). The subtype d is provisional, since only short sequences within the ORF2 are available. HEV-4 is divided into 9 subtypes (a-i) and one unclassified strain.

Strains belonging to the *Orthohepevirus B*, *C*, and *D* genera are less variable than *A* and limited number of sequences is available.

Fig. 1 Phylogenetic tree of representative members of the *Hepeviridae* family (Doceul et al., 2016)



1.1.2 Hepatitis E virus infection in humans

HEV is transmitted by the fecal-oral route and is estimated to cause every year approximately 57,000 deaths, 20 million of infections of which 3 million develop illnesses (Lozano et al., 2012). HEV infection can cause a high number of clinical manifestations from subclinical or asymptomatic to fulminant forms, with loss of liver function. The disease has a low mortality rate (2%) but reaches 20% in pregnant women. The incubation period is approximately 40 days, ranging between 2 and 10 weeks. The viremia is transitory; it occurs during the icteric phase and disappears when symptoms appear, except in the chronic form. Fecal excretion of the virus begins few days before jaundice and decreases in 2-3 weeks (Kumar et al., 2013). The acute form of hepatitis E lasts in 6-7 weeks. Jaundice is the most common symptom, can last 2-4 weeks or longer if the cholestasis is prolonged. Asthenia, fever, nausea, vomiting, abdominal pains and joints are common symptoms (Aggarwal, 2013a). The immune response, which initially increases during jaundice with a release of anti-HEV IgM, can persist in the individual from the first week of infection to 5 months. IgG antibodies are subsequently found and persist for long periods, even up to 14 years (Aggarwal, 2013b). HEV-3 and HEV-4 genotypes can lead to chronic hepatitis in organ transplant patients (e.g. liver, heart and kidney), HIV positive individuals, stem-cell-transplant patients, hematology patients receiving chemotherapy, and rheumatology patients receiving immunotherapy (Kamar et al., 2017). HEV infections (HEV-3 and HEV-4) could occur with extra-hepatic manifestations as neurological complications, Guillain-Barré syndrome, brachial neuritis, acute transverse myelitis, and acute meningoencephalitis (Kamar et al., 2010; Cheung et al., 2012).

Epidemiology in low income countries

Two epidemiological patterns are recognized: the endemic hepatitis E, causing large outbreaks in low income countries, and the zoonotic, causing sporadic cases and small outbreaks in industrialized countries (Khuroo & Khuroo, 2016b).

The endemic form is associated to HEV-1 and 2 infections, it is widespread in low income countries and it is not associated to zoonosis, no animal reservoir has been described for these genotypes. The first major epidemic was described in New Delhi in 1955 with more than 29,000 symptomatic jaundiced people and more outbreaks have been reported since then. The source of infection is associated to consumption of drinking water contaminated by human feces, caused by improper release or decontamination of sewages. Outbreaks have also been associated to poor personnel hygiene, inadequate sanitation, and an unsafe drinking water supply. In refugee camps, hepatitis E outbreaks are frequently reported. The mortality rate linked to HEV-1 and HEV-2 infections is 2% and increases up to 20% in pregnant women, if infected during the first months of pregnancy (Kamar et al., 2017). Males are more affected than females, with a ratio of 3.5:1. Symptomatic infections are

more frequent among young adults of 15-30 years old. The seroprevalence in adults in endemic areas ranges between 30% and 80%. Cases of hepatitis E in developed countries caused by HEV-1 and HEV-2 have been described in travelers returning from endemic countries (Murrison & Sherman, 2017).

Epidemiology in industrialized countries

The epidemiological pattern of HEV infections in developed countries is different, only sporadic cases and small outbreaks occur and are linked to infections by the zoonotic HEV-3 and HEV-4. Since the last decade, an increasing number of autochthonous cases has been reported in Europe (Adlhoch et al., 2016) (Fig. 2) and now the hepatitis E is considered an emerging disease. In Europe, foodborne is considered the main route of transmission, linked to the consumption of raw or undercooked meat or liver from pigs and wild boars (Pavio et al., 2015). In 2003, a small cluster of cases associated to HEV-4 infection was linked to the consumption of sashimi deer in Japan (Tei et al., 2003). Since then, several studies have reported the molecular detection of near or identical strains from human patients and consumed raw or undercooked animal (swine, wild boar, and deer) food products (liver sausages, sausages, meat, roasted piglet) (Meng et al., 1997; Meng, 2000; Okamoto et al., 2001; Smith, 2001; Teo, 2010). Infectious HEV was detected in pork liver sausage (figatelli) and the virus remained infectious up to 60°C, as proved by experimental infections in pigs, confirming a potential risk of infections linked to consumption of raw or undercooked liver or meat. People that live or work in contact with the animal reservoir of the virus as farmers, veterinarians, workers of the slaughterhouse or hunters show higher HEV seroprevalence than the general population, confirming that exposure to animal reservoirs of the zoonotic HEV can be a factor risk for humans (Pavio et al., 2017).

Milk as possible source of HEV infection in humans has been recently suggested. HEV-4 infectious strains have been detected from cow milk samples in a rural area of China (Huang et al., 2016). However, the role of milk as source of infection is still debate since other studies conducted in Germany (Baechlein & Becher, 2017) and China (Geng et al., 2018) did not reveal the presence of HEV RNA in milk from dairy cow. Nevertheless, HEV-7 (camelid HEV) infection in a transplanted patient who regularly consumed raw camel milk has been described. HEV-7 strictly related strain was also detected in camels, suggesting a possible foodborne transmission (Lee et al., 2016).

Another possible source of infection is water contaminated by human or animal feces, used for irrigation of salads, soft fruits or for harvesting of shellfish that could be a vehicle of indirect food contamination and source of HEV. However, no hepatitis E cases have been linked directly to consumption of these foods. Only epidemiological evidence supported a link between consumption

of shellfish and 4 cases of hepatitis E among passengers returning to the United Kingdom after a world cruise (Said et al., 2009).

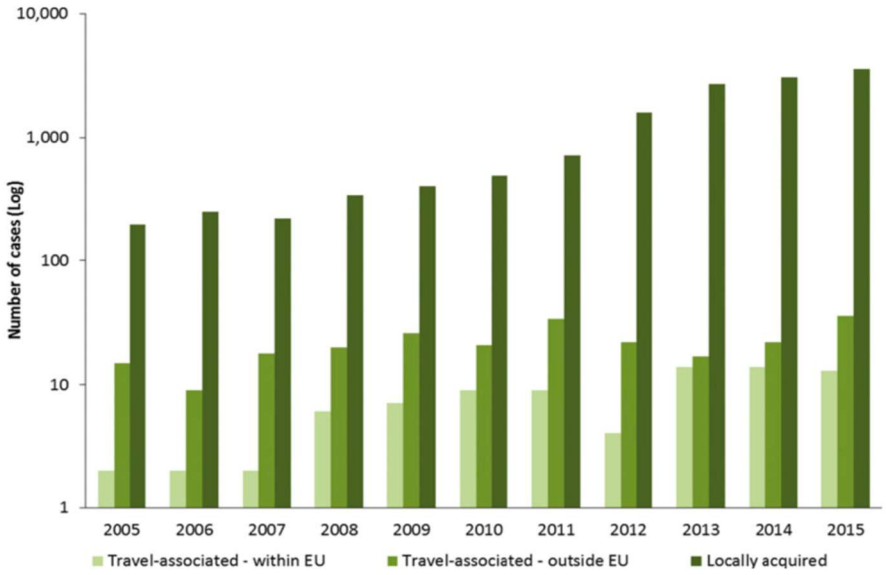
Nevertheless, in Europe several surveys have been conducted on vegetable (berries and salad) and commercial mussels, confirming the detection of HEV-3 (Diez-Valcarce et al., 2012) with a low prevalence <14% in mussels and <3.4% in fresh lettuce (Kokkinos et al., 2012; Mesquita et al., 2016). In Europe, seroprevalence data are limited to few countries where a considerable heterogeneity is shown. The seroprevalence ranges from 0.6% to 52.5% (Hartl et al., 2016). The high seroprevalence observed in some countries does not reflect the reported limited number of cases, suggesting a majority of asymptomatic infections. Furthermore, geographical differences have also been observed. In Southern France, the seroprevalence is 4 times higher than in the North of the country (Boutrouille et al., 2007), in Italy, the seroprevalence in Abruzzo region is 54%, significantly higher than in other parts of the country. The wide geographic variation in seroprevalence among different European countries and within the same country may depend on the use of different assays with different specificity, on the presence of animal reservoirs (e.g. area of intensive pig farming) or different dietary habits as consumption of raw liver sausage, traditionally eaten raw in Corsica (named figatelli), Southern France or in some Italian regions.

In Italy, a recent study reported the presence of IgG antibodies in the 8.7% of 10,011 plasma samples from blood donors, sampled throughout the Italian regions. Some regions showed higher prevalence rates such as 10-15% in Lazio, Umbria and Marche and >22% in Abruzzo and in Sardinia. The latter are considered hyperendemic areas and the observed prevalence may depend on local dietary habits, such as consumption of sausages in Abruzzo, or traditional free-range pigs farming in Sardinia (Spada et al., 2018).

In Italy, hepatitis E is a notifiable disease but until few years ago most of the cases were grouped in non-ABC hepatitis or under-diagnosed. More recently, the diagnoses of the disease are more frequent and 211 cases of acute hepatitis E were notified during the period 2007-2016 (http://www.epicentro.iss.it/problemi/epatite/EpidemiologiaItalia.asp). In Italy, between 1994 and 2009, a study was conducted on hospitalized patients affected by hepatitis E. The virus was detected in 20.6% (134/651) of tested patient sera, 16.4% of them were autochthonous cases infected by HEV-3 (Romano et al., 2011). In Europe, the number of clinical cases increased from 514 in 2005 to 5,617 cases in 2015, with 21,000 cases reported from 22 countries between 2005 and 2015. Several countries reported a considerable increase of cases which can be linked to higher awareness of clinicians, to use of better diagnostic tests and to an emerging of the disease (Ricci et al., 2017). Authocotonous infections in industrialized countries are more frequently in middle-aged men (about 60 years) with a male/female ratio of 3:1 (Mansuy et al., 2004; Bendall et al., 2008; Dalton et al.,

2008; Drobeniuc et al., 2013). Furthermore, hepatitis E is more common in individuals who consume excessive alcohol and are more prone to fatty and liver damage (Said et al., 2009; Dalton et al., 2011). HEV transmission by blood transfusion has been documented in two cases and recently the risk linked to blood-transfusion is a major concern in Europe where some additional control measures have been in place (Boxall et al., 2006; Colson et al., 2007; Matsubayashi et al., 2008).

Fig. 2 Number of human HE cases by origin of infection, 15 EU/EEA countries, 2005–2015 (Ricci et al., 2017)



1.1.3 Hepatitis E virus infection in animals

Zoonotic strains

The two HEV-3 and HEV-4 genotypes are proved as zoonotic infecting both humans and animals. Infections in animals are asymptomatic. Pigs are the main reservoirs of HEV-3 and HEV-4. In Europe, the HEV-3 is the most common in both humans and pigs. HEV-4 was firstly identified in pigs. In Asia, it is prevalent in humans while in Europe is rarely detected. HEV-4 strains were detected in Europe, probably introduced by imported animals from Asia. The other reservoirs of HEV-3 are wild boar and deer. Strains detected in rabbits are related to HEV-3 and included in a different clade by phylogenetic analyses and named HEV-3ra.

HEV in pigs

The first HEV animal strain was identified in the United States in 1995 (Meng et al., 1997). The strain showed high nucleotide and amino acids identities >80% with HEV-3 strains circulating in humans in the same geographic area, corroborating the hypothesis of zoonotic transmission. Pigs have been successfully infected intravenously with human HEV-3 and HEV-4 strains recovered from patients

with hepatitis E. However, pigs are not susceptible to infection with strains belonging to HEV-1 and/or HEV-2. Swine HEV-3 strain was used to successfully infects rhesus and chimpanzee monkeys, thus proving the transmission of the virus from pig to primates. Pigs are only infected by HEV-3 and HEV-4 which are now definitively considered zoonoses (Goel & Aggarwal, 2016). In Europe including Italy, HEV-3 is the prevalent genotype in pigs and in human cases (Zanetti et al., 1999; Di Bartolo et al., 2008; Festa et al., 2014; Lapa et al., 2015). Genotype HEV-4 mainly circulates in South East Asia (Thiry et al., 2017), and it has been rarely detected in pigs in Belgium and in Italy and in few human cases in France and Italy.

Since the first identification of HEV-3 in pigs, several studies reported the detection of HEV-3 in pig herds with prevalence up to 100% and seroprevalence ranging between 5% and 100% all over the world (Pavio et al., 2017).

The infection in pigs occurs in an early age after the loss of maternal immunity (aged 3-4 months). However, in studies conducted on animals at slaughterhouse (aged 6 months) HEV-RNA is still excreted in feces up to 18 weeks of age with the shedding peak between 3 and 8 weeks. At the end of the peak of shedding, there is the appearance of IgG (seroconversion IgM to IgG). However, the duration of acquired immunity after infection is still unknown (Ruggeri et al., 2013).

The viremia last from 1–2 weeks after infections, before shedding of the virus in feces. The main site of virus replication is liver where the highest titer of the virus was observed. HEV has been also detected in extrahepatic sites as spleen, colon and in stomach but only in experimental infected animals and with low viral load (Khuroo & Khuroo, 2016a).

Duration of HEV shedding in feces or in blood and the age of infections largely influence the persistence of the infection in farms and the risk of contaminate food. Indeed, HEV released by faces lead to accumulation of HEV in the farm environment. The repeated contacts between animals in the same pen or with contaminated environment play the main role in the HEV propagation and persistence in farm (Bouwknegt et al., 2008; de Deus et al., 2008; Walachowski et al., 2014). The duration of the immunity has never been estimated and re-infection can not be excluded. In fact, several studies reported the detection of HEV in livers from pigs at slaughter age (6-9 months) with prevalence ranging between 0.8 and 10% (Ricci et al., 2017). The presence of HEV in pig muscles was not proved since detected only sporadically (Di Bartolo et al., 2012; Feurer et al., 2018).

The foodborne transmission by ingestion of contaminated raw or undercooked pork products has been confirmed by the detection of the same strain sequences from human cases and leftover suggesting that the consumption of these products may be a risk factor for infection in humans. The first outbreak in Europe was described in France. A small cluster of cases of hepatitis E occurred in Corsica caused by direct consumption of figatelli, a raw pig liver sausages produced locally (Renou et al., 2014). Several other studies confirmed the foodborne transmission in France where the subjects reported the

consumption of raw pork products such as liver sausages (Harrison & DiCaprio, 2018). Several pork products sold at market were found to be HEV positive such as raw liver, raw sausages, liver sausages (figatelli) and patè. The percentage of detection varied a lot ranging between 1% and 47%. Raw liver sausages are the products more frequently contaminated by HEV (Yazaki et al., 2003; Feagins et al., 2007; Colson et al., 2010; Wenzel et al., 2011; Berto et al., 2012; Di Bartolo et al., 2012; Pavio et al., 2014; Di Bartolo et al., 2015; Heldt et al., 2016; Mykytczuk et al., 2017). Detection of HEV genomes does not prove the infectivity of the virus; the infectivity of HEV-3 detected in liver sausages was proved on cell culture (Berto et al., 2013). In Italy, the HEV RNA was detected in raw and dry pork liver sausages sold at market (Di Bartolo et al., 2015). One case of foodborne transmission was described in a subject, infected by HEV, who declared consumption of figatelli (liver sausages) and in the study was described the genetically correlation between the HEV human strain and those found in a restricted area of the South of France where the patient bought the figatelli (Di Bartolo et al., 2012; Garbuglia et al., 2015). In Europe and in Italy, HEV-3 is the most common genotype in pig population where the most common subtypes detected are HEV-3e and -3f (Di Bartolo et al., 2008; Festa et al., 2014; Lapa et al., 2015).

HEV in wild boar

HEV in wild boar was reported for the first time at the end of 1999 with the detection of anti-HEV antibodies (Chandler et al., 1999). The partial or complete genome sequences of wild boar HEV strains confirmed relatedness to human, deer and pig HEV-3 strains (Sonoda et al., 2004; Takahashi et al., 2014). Now wild boar is recognized as another important reservoir of the zoonotic HEV-3 genotype and is widespread in Europe and Japan (Spahr et al., 2018).

Among European countries the reported prevalence ranges between 3.7% (Caruso et al., 2015a) and 68.2% (Adlhoch et al., 2009). The difference in HEV prevalence may depend on several factors: geographical distribution, year of sampling, wild boar density, and specimen tested (Pavio et al., 2017). The seroprevalence described in wild boar in Europe (Kukielka et al., 2016) is lower than in pigs, it could also be linked to intensive pigs rearing compared to free ranging of wild boar. Another possible explanation could be a different susceptibility or clearance of the virus in pigs and wild boar. Indeed, HEV RNA has been detected in each age classes of wild boar including animals older than 2 years (Martelli et al., 2008), while in pigs the peak of infections is in animals younger than 4 months. The short length of protective immunity may lead to re-infections of older animals in wild boar. Chronic infections could occur in wild animals as observed in two naturally infected wild boars (Schlosser et al., 2014). Furthermore, in a recent study it was observed that 89% of muscle sampled from wild boar HEV positive in liver was found positive for HEV (Anheyer-Behmenburg et al., 2017).

In Italy, several studies investigated the presence of HEV in wild boar, sampled in different regions of the country, results obtained showed a wide heterogeneity of HEV circulation, with the prevalence ranging between 1.5% in Valle d'Aosta, Piemonte and Abruzzo regions and 33.5% in Latium region (liver, bile, feces). The HEV strains detected and characterized belong to HEV-3 and were classified into different clades including strains of HEV-3e, -3f, -3c and some strains that could not be classified in any subtypes (Martelli et al., 2008; Caruso et al., 2015a; Martinelli et al., 2015; Mazzei et al., 2015; Montagnaro et al., 2015; Serracca et al., 2015; Di Profio et al., 2016; Aprea et al., 2018).

The zoonotic transmission linked to consumption of raw or undercooked wild boar meat was reported in Japan and Spain. The same HEV strain, as confirmed by sequence analysis, was detected in patients affected by hepatitis E and the leftover wild boar meat consumed undercooked (Li et al., 2005a; Rivero-Juarez et al., 2017). In Italy, there is no direct evidence of cases linked to ingestion of wild boar meat, only one patient infected by HEV-3 who reported consumption of wild boar meat (Giordani et al., 2013).

The strains detected in wild boar population in Europe have been mostly classified as HEV-3e, -3f, and -3c and -3a (Vina-Rodriguez et al., 2015). Phylogenetic analysis on wild boar strains showed the evolutionary relationships (89-99% nucleotide identity) with swine and human strains.

Wild boar infected by HEV-4 has only been reported in Japan. Recently, two novel HEV classified into genotype HEV-5 and HEV-6 have been identified wild boar in Japan (Sato et al., 2011; Hara et al., 2014; Takahashi et al., 2014), but no human cases result associated to these viruses.

HEV in deer

HEV RNA and antibodies anti-HEV have been revealed in several deer species in Asia, America and Europe. Deer can be infected by HEV-3 and HEV-4 strains,closely related to wild boar, pigs and human strains of the same genotypes (Spahr et al., 2018) The first evidence of foodborne transmission of HEV was the occurrence of a small outbreak among family members infected after consumption of sashimi from Sika deer in 2003 in Japan (Tei et al., 2003; Takahashi et al., 2004; Tei et al., 2004). No cases of foodborne transmission from deer have been reported in Europe. It is not clear if deer are natural reservoirs of the infection or accidental host due to spill over event from infected wild boar or pigs.

HEV in rabbit

HEV RNA was detected in rabbits, both domestic and wild animals in Asia, Europe and USA (Spahr et al., 2018) and in one pet rabbit in Italy (Caruso et al., 2015b). The phylogenetic analysis conducted on HEV sequenced strains revealed a monophyletic clade distinct but closely related to HEV-3 and named HEV-3ra. There are no cases of hepatitis E directly linked to consumption of rabbit meat or liver but in France, 5 patients were infected with HEV strains classified as rabbit HEV-3 ra. However,

none of the patients had previously contact or consumed rabbits (Abravanel et al., 2017). The potential role of HEV-ra in human infections needs to be further investigated.

HEV in domestic ruminants

Several studies reported the presence of antibodies anti-HEV in goat, sheep and cattle. However, only in the last years HEV RNA has been detected in goats in Italy (HEV-3) (Di Martino et al., 2016) and in China (HEV-4) (Li et al., 2017; Long et al., 2017). In China, HEV-4 genotype strains were detected in sheep livers (Wu et al., 2015) and in cattle in both feces and milk (Huang et al., 2016; Yan et al., 2016). The infectivity of virus detected was proved *in vivo* in rhesus macaques which shed the virus after inoculation with HEV-contaminated raw or pasteurized milk from the infected cow (Huang et al., 2016). Very recently, HEV-4 sequences have also been detected in samples from yellow cattle in China (Yan et al., 2016).

1.1.4 HEV diagnostic techniques

Real-Time Reverse Transcription PCR

The Real-Time Reverse Transcription PCR (Real-time RT-PCR) is now considered the golden standard for the detection of the viral RNA in clinical and environmental samples. The protocol is extremely efficient and is based on a single step Real-time RT-PCR assay using a TaqMan probe annealing in a conserved region of the ORF3 (Jothikumar et al., 2006). The protocol is suitable for the 4 genotypes HEV1-4. Commercial test validated are available for human diagnoses by detection of HEV-RNA in blood.

Many laboratories also perform sequencing of HEV positive samples. Sequences analysis is used to determine the genotype, the subtype, possible mutation or novel strain and also to identify source of infection (98-100% matching between patients and animal food). Several genomic regions can be analyzed. For subtyping the most common regions is a 330bp conserved fragments within the ORF2, but other regions as RdRp and MetT are also used.

Enzyme-linked Immunosorbent Assay (ELISA)

Several commercial ELISA test are available, for both animals and humans antibodies detection. Most of ELISAs are based on capsid (ORF2) antigens with or without ORF3, since up to now only one serotype has been described the same antigen is frequently used for detection of anti-HEV in humans and animals. The ELISA specifically designed for detection of anti-HEV IgM is used as rapid bedside test for early detection of the diseases in humans.

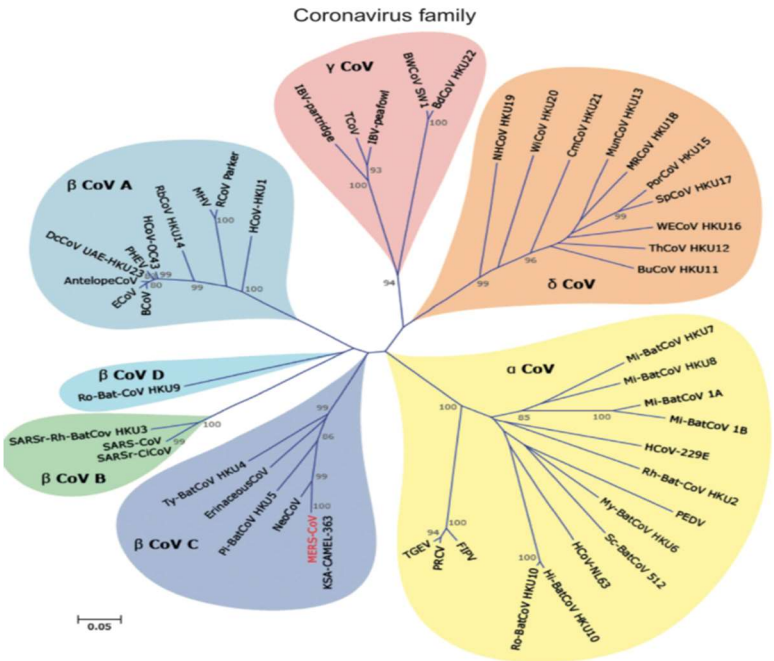
Cell culture models

The cell culture systems for diagnostic have been replaced by faster methods as quantitative real time PCR. However, it is still the best method to test the infectivity of a given sample. Besides several studies reporting HEV cultivation, growth of the virus is on cell culture is still difficult and scarcely reproducible. The human liver carcinoma (PLC/PRF/5, HepG2/C3A), cell lines derived from stem cells and human lung carcinoma cell lines (A549) are the cell lines most frequently used for HEV1-4 cultivation (Aggarwal & Goel, 2016). Clinical specimens as human serum, feces (from pig) or liver homogenates (from pig) were used as inoculum in successful growing experiments. However, several weeks (up to 50 days) are needed to obtain a moderate growth of the virus and the reproducibility is limited.

1.2 Coronavirus

Coronaviruses are the second worldwide cause of respiratory disease in human after influenza virus. Coronaviruses are classified in the *Nidovirales* order and the *Coronaviridae* family. Based on sequence comparison the family is divided into four genera: *Alpha* (*Alpha-CoV*), *Beta* (*Beta-CoV*), *Gamma* (*Gamma-CoV*) and *Deltacoronavirus* (*Delta-CoV*). The *Alpha-CoV* genus infects mammals only and is subdivided into several lineages (Fig. 3). The Alphacoronavirus 1 lineage infects pig, dog, cat, mink and ferret. The other lineages infect human, pig, bat and have been found recently in rat and shrew. The *Beta-CoV* genus infects human, mouse, calf and bat. Gamma-CoVs infect chickens and beluga whales while Delta-CoVs infect pigs and some avian species.

Fig. 3 Phylogenetic tree of 50 Coronaviruses based on RNA-dependent RNA polymerase partial nucleotide sequences (Chan et al., 2015)



1.2.1 Coronavirus features

Genome organization

Coronaviruses share a common genome organization: a 5' UTR, a large ORF1 that covers two-third of the genome, spike (S), envelope (E), membrane (M), nucleoprotein (N) and the 3' UTR, followed by a polyadenine tail. Different CoV species present also accessory proteins with important roles in viral pathogenesis (Fig. 4) (Liu et al., 2014).

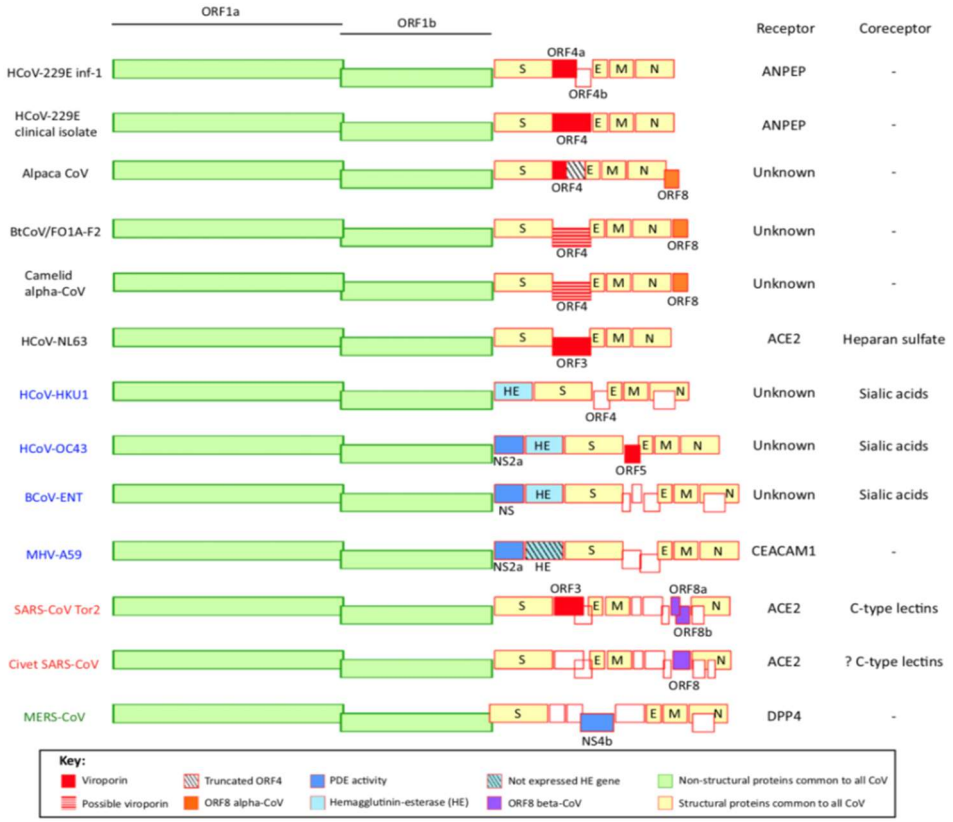
The ORF1 consists of two large over-lapping open reading frames (ORF1a and ORF1b) encoding two polyproteins: the pp1a and pp1ab. The non-structural polyproteins are processed into 16 nonstructural proteins (nsp1 to 16). The ORF1ab is characterised by a ribosomal slippage sequence (UUUAAAC), leading to the transcription of a unique large polyprotein named pp1ab (Masters, 2006). The ORF1ab produces proteins with different functions: papain-like protease (Plpro), main protease (Mpro), helices, two methyltransferase, a RNA-dependent RNA polymerase, useful to the genome replication. The function of some nsp proteins is unknown, however, several of them are antagonists of immunity response (Perlman & Netland, 2009). All the CoVs share a transcription-regulatory sequence (TRS) named leader (TRS-L) found downstream of the ORF1ab codon start. The secondary TRSs, body TRS (TRS-B), precede the other CoVs genes leading to a discontinuous transcription and synthesis of mono or policystronic subgenomic RNA. In fact, the main feature of *Nidovirales* order derives from the nested 3′ mRNAs as *nido* translated from Latin as "nest" (Perlman & Netland, 2009).

Among RNA viruses, CoVs acquire exceptionally long genomes. The hypothesis is that *Nidovirales* viruses acquired enzymes that increase the fidelity of RNA replication. These enzymes are encoded by ORF1ab and show RNA 3'-to-5' exoribonuclease (ExoN) and an endoribonuclease (NendoU) (Lauber et al., 2013). The genome expansion allowed the acquisition of novel genes encoding accessory proteins that may promote virus adaptation or be antagonists of immune responses, increasing the virulence.

Among structural proteins, there is the spike protein that forms the characteristic spherical virion with club-shape spike projections and gives them the appearance of a solar corona, prompting the name coronaviruses (Barcena et al., 2009). The spike protein (~150 kDa) has two distinct subunits: the S1 and S2 domain. The S1 is divided into N- terminal domain (NTD) and the receptor-binding domain (RBD, also referred to as C-terminal domain or CTD). The S2 includes the fusion peptide, two heptad repeats (HR1 and HR2), and the transmembrane region (Graham & Baric, 2010). The S glycoprotein is a class I fusion protein and is able to bind the host receptor. The S protein is recognized by host membrane cell protease and separated into S1 and S2 polypeptides (Bosch et al., 2003). The M protein gives the virion its shape, exists as dimer, promotes the membrane curvature and binds the

nucleocapsid. The E protein is involved into the assembly and release of the virus. The N protein is a helically symmetrical protein phosphorylated capable of binding RNA *in vitro*. The protein is involved in the encapsidation of the viral genome into viral particles, identifying the TRSs and the genomic packaging signal the nsp3 and M protein (Fehr & Perlman, 2015).

Fig. 4 CoV genome organization. Information on receptor and co-receptor usage is reported (Forni et al., 2015)



Coronavirus replication cycle

CoVs infection starts with the binding of S1 spike protein domain to the cell receptor triggering a conformational change that promotes membrane fusion of the virus and the cell through the S2 domain. Different CoV species recognize different host receptors. Human CoVs (HCoV), the HCoV-229E, binds the aminopeptidase N, HCoV-OC43 and HCoV-HKU1 the 9-O-acetylated sialic acid, the SARS-CoV the angiotensin-converting enzyme 2 (ACE2), the MERS-CoV the dipeptidyl peptidase 4 (DPP4) (Fehr & Perlman, 2015). Other routes are recognised to support the membrane fusion and the CoV cell entry. The transmembrane protease serine 2 (TMPRSS2) and trypsin-like protease TMPRSS11D cleavage the S1/S2 domains for non-endosomal virus entry of HCoV-229E and SARS-CoV or the furin, a serine endopeptidase, for MERS-CoV (Bertram et al., 2011; Bertram et al., 2013; Millet & Whittaker, 2014). Some studies have described the role of the low pH in the

cell and the endosomal cysteine protease cathepsins, helping to CoV cell entry as cathepsin L in SARS-CoV and MERS-CoV entry (Simmons et al., 2005; Bosch et al., 2008; Qian et al., 2013).
The replication phase begins with the translation of ORF1 into pp1a and pp1b polyproteins. RNA secondary structures cause the ribosome block on the slippery sequence. The -1 frameshift is caused by the moving back of one nucleotide of the ribosome that extends the translation into pp1ab. The pp1ab contains the nsp 1-16, cleaved individually. Among nsps, there are the papain-like proteases (PLpro) and the main protease, or Mpro. The PLpro cleaves between nsp1/2, nsp2/3 and nsp3/4 while the Mpro cleaves between the remaining 11 cleavage sites. After the translation of RNA-dependent RNA polymerase (RdRp), the RNA helicase domain and RNA 5′-triphosphatase activity, the exoribonuclease (ExoN) fidelity activity and N7-methyltransferase activity, nsp16 with 2′-O-methyltransferase activity, the CoV positive genome is transcribed into a complete negative strand template and negative-stranded subgenomic mRNAs. Then, the subgenomic mRNA are transcribed and translated to encode the structural and accessory proteins (Lim et al., 2016). The hypothesis is that the RdRp pauses on the TRS sequences. During the pause the RdRp can continue the elongation or move back to amplify from the leader sequence at the 5′ end, guided by complementarity of the TRS to the leader TRS (TRS-L) (Sethna et al., 1991). The viral structure proteins: S, E and M are translated and inserted into the endoplasmic reticulum (ER) and move to the Golgi intermediate compartment (ERGIC). The M protein directs the virus-like particles (VLPs) formation, along with E, to produce coronavirus envelopes. The M protein interacts with N protein to complete the virion assembly. The N protein encapsides the viral genomes into viral structural proteins contained in the ERGIC to form mature virions. The S protein is not essential for assembly; however, it is incorporated into virions after its interaction with M protein. Following the assembly, the virions are transported in vescicles to the cell surface and released by exocytosis (Lim et al., 2016).

1.2.2 Alpha-CoVs and Beta-CoVs

Among Coronavirus, humans are infected by two genera only: Alpha-CoVs and Beta-CoVs. Two Alpha-CoVs (HCoV-229E, HCoV-NL63) and two Beta-CoVs (HCoV-OC43, HCoV-HKU1) cause mild respiratory disease where only inter-human transmission is possible. Among Alpha-CoVs and Beta-CoVs, approximately 20 CoV species cause mild to severe disease in livestock and companion animals with several economic losses to commercial activities. However, only few species show evolutionary correlation with human strains and none has showed a zoonotic behaviour. Among Beta-CoVs, in the last decade, two strains showed a zoonotic potential causing severe respiratory diseases leading to more than 8000 cases and 1000 deaths worldwide: the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and the Middle East Respiratory Disease (MERS-CoV). Phylogenetic

studies on human and animal strains have showed phylogenetic linkages between human and bat strains, considered the main reservoir and at the origin of coronaviruses to date.

Human Coronaviruses

The two Alpha-CoVs, HCoV-229E and HCoV-NL63, and two Beta-CoVs, HCoV-OC43 and HCoV-HKU1, are endemic in human population and cause mild respiratory disease. They are characterized by an inter-human transmission by personal contact, coughing, sneezing or contact with contaminated surface. HCoV-NL63 and HCoV-HKU1 cause the 15-30% of the respiratory infections every year and were recently described after the SARS-CoV outbreak in 2002. They have been detected worldwide and most of the infections are reported during the cold seasons causing a common cold. Severe disease is reported in children under the age of 6, the elderly, and in individuals with pre-existing medical conditions. Usually CoV infections cause cough, fever, sore throat, rhinitis, expectoration, and lower respiratory tract infection, such as bronchitis, bronchiolitis or pneumonia. The HCoV-229E and the HCoV-OC43 were the first human CoVs described and identified 50 years ago. They are responsible of mild respiratory tract infections such as the common cold (Su et al., 2016). The HCoV-229E strains show worldwide lower sequence divergence then the HCoV-OC43, which has showed significant genetic variability. In Italy, few studies only have described the circulation of these strains (Gerna et al., 2006; Bosis et al., 2007; Minosse et al., 2008) .

Zoonotic Coronaviruses

Over the past 10 years, two highly pathogenic coronaviruses have been reported worldwide, causing severe respiratory disease in human with thousands of cases of infections and deaths: the Severe Acute Respiratory Syndrome and the Middle East Respiratory Virus coronaviruses. These two CoV species are the only ones that have showed a zoonotic behaviour. The zoonotic transmission usually occurs by a spillover event from an intermediate amplifier host, as an animal host, to human. The surveillance studies and phylogenetic analysis on human and animals CoV strains have identified the palm civets for SARS-CoV and dromedary camels for MERS-CoV as intermediate host.

SARS-CoV

The Severe Acute Respiratory Syndrome is a viral disease of human, caused by the SARS Coronavirus. The first case was identified in 2002 in the Guangdong Province of China and the last in December 2003 in the same province. The WHO reported 37 countries involved around the world, more than 8000 cases and 775 deaths resulting in a mortality rate of 10% (http://www.who.int/csr/sars/en/). The incubation period ranges from 2 to 10 days. The initial symptoms are fever, muscle pain, migraine, cough, slight respiratory problem or diarrhea with active virus shedding. One-third of the infected patients recover, while two-thirds develop pneumonia with

breathing difficulties, requiring mechanical respirator. In fact, the most common cause of death is the respiratory failure. Other causes of death are cardiac and liver failure (Vijayanand et al., 2004).

The inter-human transmission can occur by respiratory aerosols and by touching contaminated surfaces or the mouth, nose or eye of infected individuals after the onset of illness. The SARS-CoV base-case reproduction number (R0) ranges between 2 and 4, excluding the Super Spreading Events (SSE). The SSE patients can excrete high titers of virus leading to more than 300 infections from a single patient. To contain the virus spread, different recommendations have been applied such as patient isolation, quarantine and hand washing. The application of these procedures has resulted in a significant decrease of the R0 values (<1) (http://www.who.int/csr/sars/en/WHOconsensus.pdf).

This pathogen has animal origin. Bats are considered the main reservoir while the masked palm civet acted as amplification host causing infection in human.

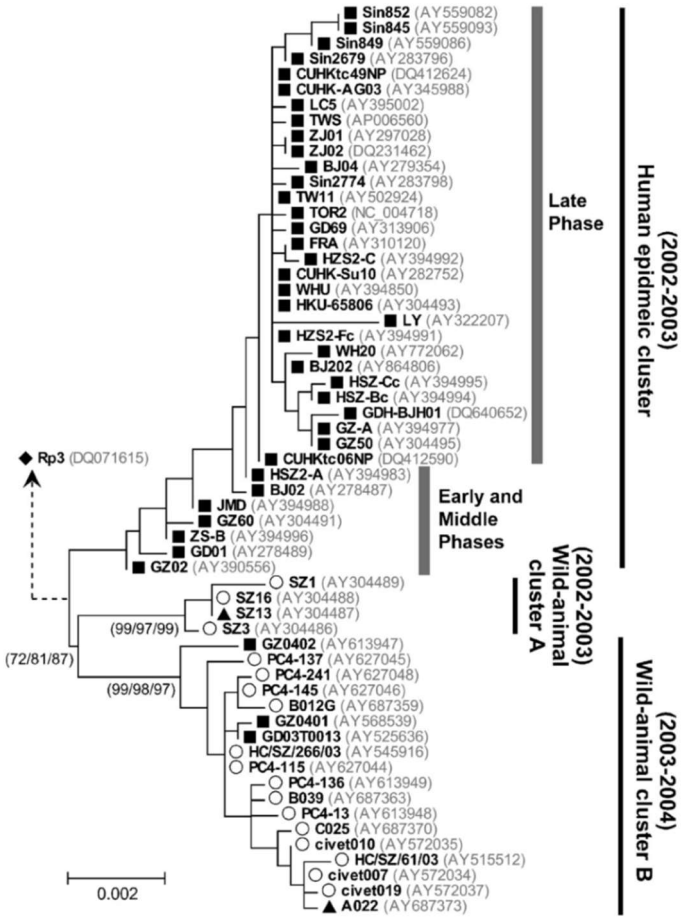
The SARS-CoVs epidemic is divided into three phases (Fig. 5). The early phase refers to the first documented infection and suddenly spread of the viruses in Guangdong province in 2002. Phylogenetic analysis on strains obtained from the early phase formed multiple clusters, representing the independent outbreaks caused by the multiple zoonotic transmission events (Song et al., 2005; Wang et al., 2005b) . The palm civets strains shared >99.0% nucleotide sequence identity at full genome level with human SARS-CoV strains of the early phase, confirming the civets as source of the early phase infections (Guan et al., 2003; Kan et al., 2005). The evidences of animal origin were supported when SARS-CoV positive palm civets strains were found in Chinese animal markets (Guan et al., 2003).

The middle phase refers to the first documented SSE, where a nephrologist was infected in Guangdong province at the end of 2003 and after a travel in Hong Kong was hospitalized starting several outbreaks. The strains detected during the middle phase formed monophyletic cluster including the Hong Kong's first patient and the ones detected in Vietnam, Singapore and Canada (Guan et al., 2004). The strains detected during the early and middle phases formed distinct clusters, showing lower correlation between them.

The late phase refers to several independent outbreaks begun in 2003 and lasted until 2004 in Guangdong province. The strains detected during this phase formed several distinct clusters and were related to civet strains confirming the direct zoonotic transmission events (Song et al., 2005; Lam et al., 2008). The strains detected in civets during the 2002/2003 and 2003/2004 outbreaks showed low genetic relatedness and formed distinct cluster suggesting that the two outbreak events were independent (Song et al., 2005; Wang et al., 2005a). However, the palm civets have never been considered as the natural reservoir of the virus due to the low genetic variability of SARS-CoV strains and their relatedness to the human strains suggesting the acquisition of the virus shortly before the transmission to human (Song et al., 2005; Lam et al., 2008). In addition, natural reservoirs usually do

not develop signs or symptoms from infection while civets experimentally infected with SARS-CoV developed symptoms such as fever, lethargy, and loss of aggressiveness (Wu et al., 2005). SARS-CoV strains, named SARS-like CoVs, were detected from the genus *Rhinolophus* (horseshoe bats) in China, sharing approximately 90% nucleotide identity to those detected in human and civets and leading to the hypothesis that bats are the natural reservoir of SARS-CoVs (Lau et al., 2005; Li et al., 2005b). However, how bat strains were transmitted to civets is still unknown. In addition, if bats transmitted the virus to civets shortly before the human emergence, these strains should be almost identical. On the contrary CoVs bat and civet strains have showed only distant phylogenetic relationship and the direct ancestor of human and civet strains has never been found (Hon et al., 2008).

Fig. 5 A phylogenetic tree based on spike gene nucleotide sequences from SARS-CoV isolated from humans, civets and raccoon dogs. Sequences from humans, civets, raccoon dogs and bats were indicated with symbols square, circle, triangle and rhombus, respectively (Yip et al., 2009)



MERS-CoV

Middle east respiratory syndrome coronavirus (MERS-CoV) was identified for the first time in 2012 in the Kingdom of Saudi Arabia (KSA) (Zaki et al., 2012). Since then, 2,248 confirmed cases have

been reported from 27 countries with 791 deaths. The majority of MERS-CoV infections has been restricted to the Arabian Peninsula (83% of the cases), however, a smaller number of cases has been reported in people who have traveled to the Arabian Peninsula or have been in contact with infected people. Cases of infections are still reported. In 2018, in KSA 110 cases were reported, 33 of them were fatal (WHO, http://www.who.int/emergencies/mers-cov/en/).

The largest outbreak described outside of KSA has been reported in Korea and started from a patient who traveled in endemic area in 2015. In 2 months, 186 cases were confirmed and 36 patients succumbed. The MERS-CoV transmission was amplified in nosocomial settings (Cowling et al., 2015). Excluding the outbreaks reported in KSA and Korea, several sporadic cases have been reported worldwide, including Italy, with one described case only and no death.

The incubation period ranges between 2 and 14 days with no developing symptoms. The disease presents a wide range diversity of clinical manifestation: from asymptomatic to mild respiratory symptoms as fever, cough, and myalgia to severe pneumonitis and respiratory failure (Drosten et al., 2013). The main risk factors of infections are the close contacts with dromedary camels and health workers who have worked in contact with MERS patients. In addition, higher risks were reported for male gender, chronic disease and immunocompromised status, in fact, healthy infected adults showed only mild or subclinical disease (Drosten et al., 2013). Strains isolated from single outbreaks are almost identical, suggesting that the virus is not adapting to human populations and people infected at different times and places showed strains with low number of mutations (Cotten et al., 2014). The R0 value of MERS-CoV is <1, however, it increases in health care settings (>1) and decreases in the same conditions, when infection prevention applications and control measures as patient isolation are applied. Sequence analysis of MERS-CoV variants showed that, during the early phase in KSA, the positive selection targeted regions useful for virus cell entry, excluding the RDB region, affecting the host or tissue tropism (Cotten et al., 2014; Forni et al., 2015). In a second moment, during the South Korean outbreak, where the virus spread rapidly, point mutations in the RBD were found. These mutated regions contained immune epitopes and evolved to avoid the binding of neutralizing antibodies but decreased the binding to the cellular receptor, with the emergence of variants able to escape immune responses (Kim et al., 2016).

As described above dromedary camels are considered the source of zoonotic transmission. After the MERS-CoV outbreaks, different mammal species were tested for anti-MERS-CoV antibodies as camels, goats, horses, chickens, sheep, and poultry. Only dromedary camels resulted positive to the detection. In addition, serum samples collected in Africa in 1982 and in Saudi Arabia in 1992 resulted positive, suggesting the circulation of the virus in dromedaries for over 20 years. Dromedary camels from Australia, Canada, the United States of America, Germany, the Netherlands or Japan resulted negative to anti-MERS-CoV antibodies detection. MERS-CoV was detected in oro-nasal and fecal

samples from dromedary camels in multiple locations in the Arabian Peninsula suggesting that direct contact with a sick animal may lead to human transmission (Omrani et al., 2015). One study reported the high seropositivity for MERS-CoV among slaughterhouse workers, while several studies reported no infection in occupational exposure, indicating that the transmission animal-human is inefficient (Hemida et al., 2015; Muller et al., 2015). In the Middle East regions, the urbanization has concentrated the camel farms next to the main cities. In addition, camels are source of milk and meat; the consumption of unpasteurized milk is common and the camel urine is widely used as medicine. Interestingly, MERS-CoV RNA has been reported in raw milk collected in a marketplace in Qatar (Reusken et al., 2014). These findings increase the opportunity of human to come in contact with infected camels (Ying et al., 2014; Luke et al., 2016; Zumla et al., 2016). Phylogenetic analysis on strains detected in a herd of dromedary camels in Dubai were almost identical to those circulating in Eastern Saudi Arabia, indicating that animal movement may introduce virus into herds (Wernery, 2014). The strains collected from human cases and camels in Qatar were identical as well. People that live and work in contact with camels had seroprevalences (~3.6–6.4%) higher than the general population (0.15%) (Reusken et al., 2016).

Human and camel MERS-CoV strains are mainly related to two bat coronaviruses HKU4 and HKU5 (van Boheemen et al., 2012). Experimental *in vitro* has showed that HKU4 and MERS-CoV can bind to human and bat DPP4 in order to enter cells (Wang et al., 2014; Yang et al., 2014). Closely related coronavirus from bats were obtained from several bat species worldwide (Annan et al., 2013; Anthony et al., 2013; Ithete et al., 2013; Lelli et al., 2013; Corman et al., 2014; Yang et al., 2014). A short fragment of the RdRp region was amplified from the RNA extracted from the fecal sample, obtained from an Egyptian tomb bat (*Taphozous perforatus*) in Bisha, South Arabia (Memish et al., 2013). Most of the studies are based on the sequencing of short CoVs genomic regions strains, detected from African and Chinese bats, which were fully sequenced and showed correlation to MERS-CoVs and considered as MERS-CoV like strain (Corman et al., 2014; Yang et al., 2014; Anthony et al., 2017). These findings suggested the origin of MERS-CoVs in bats. Despite the relatedness of MERS-CoV with bat strains, MERS-CoV has never been isolated from bats. In addition, the contact between human and bats seems to be rare. Indeed, bats are not considered the source of infection and transmission to human but the main reservoir.

Bat Coronaviruses

Bats, classified in the Chiroptera order, are represented by more than 1,200 species found worldwide and are the second mammalian order with majority of species after rodents. Bats are the reservoir of several known and unknown zoonotic viruses. Several features of bats allow them to be considered the best reservoir. In particular the flight ability, the longevity, the large colonies allow them to

acquire pathogens and spread them over a wide geographical range, increasing their interactions with humans or livestock animals and increasing the risk of interspecies or intraspecies transmission. Indeed, bats have been identified as reservoir for several emerging zoonotic like Hendra virus, Nipah virus, Ebola and Marburg viruses. Coronaviruses present a large diversity of strains in bats more than in other mammalian hosts. Bat CoV strains detected to date are classified in the genera Alpha-CoVs or Beta-CoVs. Since the SARS outbreak in 2002, the CoV surveillance in bats has been increased worldwide and the number of CoV sequences available online has increased exponentially. However, many studies are based on different PCR assays targeting different regions within the RdRp and obtaining amplicons of different length from approximately 100 bases to the complete genome. The short sequences complicate the phylogenetic analysis and may not show the real relationships between CoV species. A first classification criterion was established using the pairwise amino acid distances on a translated 816 nucleotide RdRp (nsp12) fragment using as threshold value 4.8% for Alpha-CoVs and 6.3% for Beta-CoVs (Drexler et al., 2010; Tao et al., 2012). However, with the increasing of the number of sequences and the genetic CoVs diversity this method lost his discriminatory power. Few studies only retrieved complete genome from bat CoV strains allowing the complete characterisation of the strains. In addition, sometimes the lack of complete related genomes precludes a complete characterization. The International Committee for the Taxonomy of Viruses (ICTV) has provided a classification criteria when a complete genome is available, using an amino acid sequence identity threshold value of 90% in the seven conserved replicase concatenated domains: nsp3 (ADRP), nsp5 (3CLpro), nsp12 (RdRp), nsp13 (He11), nsp14 (ExoN), nsp15 (NendoU) and nsp16 (O-MT) (de Groot 2012). Among Alpha and Beta-CoVs, 15 species are recognized to date (Drexler 2014). The Alpha-CoVs are classified into 11 species; 6 of them have been detected in bats: *Miniopterus bat coronavirus 1*, *Bat coronavirus CDPHE15*, *Miniopterus bat coronavirus HKU8*, *Rhinolophus bat coronavirus HKU2*, *Bat coronavirus HKU10*, *Scotophilus bat coronavirus 512*. Beta-CoVs are classified into 10 species; 3 of them have been detected only in bats: *HKU4*, *HKU5*, *HKU9*. However, some strains are not assigned to any species to date, underlining the high heterogeneity of CoVs in bats.

Surveillance and phylogenetic studies have detected a large number of CoVs species in bats, emphasizing the association of CoV species to single host genera independently from sampling location (Cui et al., 2007; Vijaykrishna et al., 2007). These findings may depend on the capacity of the virus to infect closely related hosts only (de Vienne et al., 2013). For example, SARS-like CoVs have been detected in Rhinolophus bat species in Asia, Europe and Africa (Drexler et al., 2014; Bourgarel et al., 2018; Gouilh et al., 2018). A recent phylogenetic study, based on a short region within the RdRp of bat CoV sequences available online, has confirmed that distinct CoV species are

associated to different bat genera. In addition, one CoV species may be associated to more than one bat species that belong to the same bat genera or that co-roost (Leopardi et al., 2018).

Besides the strains detected in bat only, some strains have showed relatedness to human CoVs species. Not only SARS-CoV and MERS-CoV showed relatedness with bat strains, but also HCoV-229E and HCoV-NL63 strains showed relatives with American tricolored bat (*Perimyotis subflavus*) and Kenyan *Triaenops afer* species, suggesting that those human strains may have bats as potential reservoir (Pfefferle et al., 2009; Huynh et al., 2012; Corman et al., 2015). In Europe the presence of CoVs has been described in more than 20 bat species from Germany, Spain, Luxembourg, The Netherlands, United kingdom, France and Hungary (Gloza-Rausch et al., 2008; Reusken et al., 2010; August et al., 2012; Kemenesi et al., 2014; Goffard et al., 2015; Monchatre-Leroy et al., 2017; Pauly et al., 2017). In Italy, three studies have detected CoVs in bats, mostly from North Italy, in *Myotis nattereri, Myotis daubentonii, Myotis myotis, Rhinolophus hipposideros, Hypsugo savii, Pipistrellus kuhlii, Pipistrellus pipistrellus, Nyctalus noctula, Epseticus serotinus, Myotis blythii, Myotis oxygnathus and Plecotus auritus* species (Lelli et al., 2013; De Benedictis et al., 2014; Rizzo et al., 2017).

1.3 References

Abravanel, F., Lhomme, S., El Costa, H., Schvartz, B., Peron, J.M., Kamar, N., et al., 2017: Rabbit Hepatitis E Virus Infections in Humans, France. *Emerging Infectious Diseases,* **23,** 1191-1193.

Adlhoch, C., Avellon, A., Baylis, S.A., Ciccaglione, A.R., Couturier, E., de Sousa, R., et al., 2016: Hepatitis E virus: Assessment of the epidemiological situation in humans in Europe, 2014/15. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology,* **82,** 9-16.

Adlhoch, C., Wolf, A., Meisel, H., Kaiser, M., Ellerbrok, H. and Pauli, G., 2009: High HEV presence in four different wild boar populations in East and West Germany. *Veterinary microbiology,* **139,** 270-278.

Aggarwal, R., 2013a: Hepatitis E: clinical presentation in disease-endemic areas and diagnosis. *Seminars in Liver Disease,* **33,** 30-40.

Aggarwal, R., 2013b: Hepatitis e: epidemiology and natural history. *Journal of clinical and experimental hepatology,* **3,** 125-133.

Aggarwal, R. and Goel, A., 2016: Advances in hepatitis E - I: virology, pathogenesis and diagnosis. *Expert review of gastroenterology & hepatology,* **10,** 1053-1063.

Anheyer-Behmenburg, H.E., Szabo, K., Schotte, U., Binder, A., Klein, G. and Johne, R., 2017: Hepatitis E Virus in Wild Boars and Spillover Infection in Red and Roe Deer, Germany, 2013-2015. *Emerging infectious diseases,* **23,** 130-133.

Annan, A., Baldwin, H.J., Corman, V.M., Klose, S.M., Owusu, M., Nkrumah, E.E., et al., 2013: Human Betacoronavirus 2c EMC/2012-related Viruses in Bats, Ghana and Europe. *Emerging infectious diseases,* **19,** 456-459.

Anthony, S.J., Gilardi, K., Menachery, V.D., Goldstein, T., Ssebide, B., Mbabazi, R., et al., 2017: Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus. *mBio,* **8**.

Anthony, S.J., Ojeda-Flores, R., Rico-Chavez, O., Navarrete-Macias, I., Zambrana-Torrelio, C.M., Rostal, M.K., et al., 2013: Coronaviruses in bats from Mexico. *The Journal of General Virology,* **94,** 1028-1038.

Aprea, G., Amoroso, M.G., Di Bartolo, I., D'Alessio, N., Di Sabatino, D., Boni, A., et al., 2018: Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy. *Transboundary and emerging diseases,* **65,** e25-e31.

August, T.A., Mathews, F. and Nunn, M.A., 2012: Alphacoronavirus Detected in Bats in the United Kingdom. *Vector-Borne and Zoonotic Diseases,* **12,** 530-533.

Baechlein, C. and Becher, P., 2017: No Evidence for Zoonotic Hepatitis E Virus Infection Through Dairy Milk in Germany. *Hepatology,* **65,** 394-395.

Balayan, M.S., Andjaparidze, A.G., Savinskaya, S.S., Ketiladze, E.S., Braginsky, D.M., Savinov, A.P., et al., 1983: Evidence for a virus in non-A, non-B hepatitis transmitted via the fecal-oral route. *Intervirology,* **20,** 23-31.

Barcena, M., Oostergetel, G.T., Bartelink, W., Faas, F.G.A., Verkleij, A., Rottier, P.J.M., et al., 2009: Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. *Proceedings of the National Academy of Sciences of the United States of America,* **106,** 582-587.

Barzon, L., Lavezzo, E., Militello, V., Toppo, S. and Palu, G., 2011: Applications of Next-Generation Sequencing Technologies to Diagnostic Virology. *International journal of molecular sciences,* **12,** 7861-7884.

Bendall, R., Ellis, V., Ijaz, S., Thurairajah, P. and Dalton, H.R., 2008: Serological response to hepatitis E virus genotype 3 infection: IgG quantitation, avidity, and IgM response. *Journal of medical virology,* **80,** 95-101.

Berto, A., Grierson, S., Hakze-van der Honing, R., Martelli, F., Johne, R., Reetz, J., et al., 2013: Hepatitis E Virus in Pork Liver Sausage, France. *Emerging infectious diseases,* **19,** 264-266.

Berto, A., Martelli, F., Grierson, S. and Banks, M., 2012: Hepatitis E Virus in Pork Food Chain, United Kingdom, 2009-2010. *Emerging infectious diseases,* **18,** 1358-1360.

Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., et al., 2013: TMPRSS2 Activates the Human Coronavirus 229E for Cathepsin-Independent Host Cell Entry and Is Expressed in Viral Target Cells in the Respiratory Epithelium. *Journal of Virology,* **87,** 6150-6160.

Bertram, S., Glowacka, I., Muller, M.A., Lavender, H., Gnirss, K., Nehlmeier, I., et al., 2011: Cleavage and Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by Human Airway Trypsin-Like Protease. *Journal of Virology,* **85,** 13363-13372.

Bexfield, N. and Kellam, P., 2011: Metagenomics and the molecular identification of novel viruses. *Veterinary Journal,* **190,** 191-198.

Bosch, B.J., Bartelink, W. and Rottier, P.J.M., 2008: Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide. *Journal of Virology,* **82,** 8887-8890.

Bosch, B.J., van der Zee, R., de Haan, C.A.M. and Rottier, P.J.M., 2003: The coronavirus spike protein is a class I virus fusion protein: Structural and functional characterization of the fusion core complex. *Journal of Virology,* **77,** 8801-8811.

Bosis, S., Esposito, S., Niesters, H.G.M., Tremolati, E., Pas, S., Principi, N., et al., 2007: Coronavirus HKU1 in an Italian pre-term infant with bronchiolitis. *Journal of Clinical Virology,* **38,** 251-253.

Bourgarel, M., Pfukenyi, D.M., Boue, V., Talignani, L., Chiweshe, N., Diop, F., et al., 2018: Circulation of Alphacoronavirus, Betacoronavirus and Paramyxovirus in Hipposideros bat species in Zimbabwe. *Infection Genetics and Evolution,* **58,** 253-257.

Boutrouille, A., Bakkali-Kassimi, L., Cruciere, C. and Pavio, N., 2007: Prevalence of anti-hepatitis E virus antibodies in French blood donors. *Journal of clinical microbiology,* **45,** 2009-2010.

Bouwknegt, M., Engel, B., Herremans, M.M.P.T., Widdowson, M.A., Worm, H.C., Koopmans, M.P.G., et al., 2008: Bayesian estimation of hepatitis E virus seroprevalence for populations

with different exposure levels to swine in The Netherlands. *Epidemiology and infection,* **136,** 567-576.

Boxall, E., Herborn, A., Kochethu, G., Pratt, G., Adams, D., Ijaz, S., et al., 2006: Transfusion-transmitted hepatitis E in a 'nonhyperendemic' country. *Transfusion Medicine,* **16,** 79-83.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., et al., 2000: Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nature biotechnology,* **18,** 630-634.

Cao, D. and Meng, X.J., 2012: Molecular biology and replication of hepatitis E virus. *Emerging microbes & infections,* **1,** e17.

Caruso, C., Modesto, P., Bertolini, S., Peletto, S., Acutis, P.L., Dondo, A., et al., 2015a: Serological and virological survey of hepatitis E virus in wild boar populations in northwestern Italy: detection of HEV subtypes 3e and 3f. *Archives of virology,* **160,** 153-160.

Caruso, C., Modesto, P., Prato, R., Scaglione, F.E., De Marco, L., Bollo, E., et al., 2015b: Hepatitis E Virus: First Description in a Pet House Rabbit. A New Transmission Route for Human? *Transboundary and emerging diseases,* **62,** 229-232.

Chan, J.F., Lau, S.K., To, K.K., Cheng, V.C., Woo, P.C. and Yuen, K.Y., 2015: Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. *Clinical microbiology reviews,* **28,** 465-522.

Chandler, J.D., Riddell, M.A., Li, F., Love, R.J. and Anderson, D.A., 1999: Serological evidence for swine hepatitis E virus infection in Australian pig herds. *Veterinary microbiology,* **68,** 95-105.

Cheung, M.C., Maguire, J., Carey, I., Wendon, J. and Agarwal, K., 2012: Review of the neurological manifestations of hepatitis E infection. *Annals of hepatology,* **11,** 618-622.

Colson, P., Borentain, P., Queyriaux, B., Kaba, M., Moal, V., Gallian, P., et al., 2010: Pig Liver Sausage as a Source of Hepatitis E Virus Transmission to Humans. *Journal of Infectious Diseases,* **202,** 825-834.

Colson, P., Coze, C., Gallian, P., Henry, M., De Micco, P. and Tamalet, C., 2007: Transfusion-associated hepatitis E, France. *Emerging infectious diseases,* **13,** 648-649.

Corman, V.M., Baldwin, H.J., Tateno, A.F., Zerbinati, R.M., Annan, A., Owusu, M., et al., 2015: Evidence for an Ancestral Association of Human Coronavirus 229E with Bats. *Journal of Virology,* **89,** 11858-11870.

Corman, V.M., Ithete, N.L., Richards, L.R., Schoeman, M.C., Preiser, W., Drosten, C., et al., 2014: Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *Journal of Virology,* **88,** 11297-11303.

Cotten, M., Watson, S.J., Zumla, A.I., Makhdoom, H.Q., Palser, A.L., Ong, S.H., et al., 2014: Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *mBio,* **5**.

Cowling, B.J., Park, M., Fang, V.J., Wu, P., Leung, G.M. and Wu, J.T., 2015: Preliminary epidemiological assessment of MERS-CoV outbreak in South Korea, May to June 2015. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin,* **20,** 7-13.

Cui, J., Han, N.I.J., Streicker, D., Li, G., Tang, X.C., Shi, Z.L., et al., 2007: Evolutionary relationships between bat coronaviruses and their hosts. *Emerging infectious diseases,* **13,** 1526-1532.

Dalton, H.R., Bendall, R., Ijaz, S. and Banks, M., 2008: Hepatitis E: an emerging infection in developed countries. *Lancet Infectious Diseases,* **8,** 698-709.

Dalton, H.R., Bendall, R.P., Rashid, M., Ellis, V., Ali, R., Ramnarace, R., et al., 2011: Host risk factors and autochthonous hepatitis E infection. *European journal of gastroenterology & hepatology,* **23,** 1200-1205.

De Benedictis, P., Marciano, S., Scaravelli, D., Priori, P., Zecchin, B., Capua, I., et al., 2014: Alpha and lineage C betaCoV infections in Italian bats. *Virus genes,* **48,** 366-371.

de Deus, N., Casas, M., Peralta, B., Nofrarias, M., Pina, S., Martin, M., et al., 2008: Hepatitis E virus infection dynamics and organic distribution in naturally infected pigs in a farrow-to-finish farm. *Veterinary microbiology,* **132,** 19-28.

de Vienne, D.M., Refregier, G., Lopez-Villavicencio, M., Tellier, A., Hood, M.E. and Giraud, T., 2013: Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytologist,* **198,** 347-385.

Di Bartolo, I., Angeloni, G., Ponterio, E., Ostanello, F. and Ruggeri, F.M., 2015: Detection of hepatitis E virus in pork liver sausages. *International journal of food microbiology,* **193,** 29-33.

Di Bartolo, I., Diez-Valcarce, M., Vasickova, P., Kralik, P., Hernandez, M., Angeloni, G., et al., 2012: Hepatitis E virus in pork production chain in Czech Republic, Italy, and Spain, 2010. *Emerging infectious diseases,* **18,** 1282-1289.

Di Bartolo, I., Martelli, F., Inglese, N., Pourshaban, M., Caprioli, A., Ostanello, F., et al., 2008: Widespread diffusion of genotype 3 hepatitis E virus among farming swine in Northern Italy. *Veterinary microbiology,* **132,** 47-55.

Di Martino, B., Di Profio, F., Melegari, I., Sarchese, V., Robetto, S., Marsilio, F., et al., 2016: Detection of hepatitis E virus (HEV) in goats. *Virus Research,* **225,** 69-72.

Di Profio, F., Melegari, I., Sarchese, V., Robetto, S., Marruchella, G., Bona, M.C., et al., 2016: Detection and genetic characterization of hepatitis E virus (HEV) genotype 3 subtype c in wild boars in Italy. *Archives of virology,* **161,** 2829-2834.

Diez-Valcarce, M., Kokkinos, P., Soderberg, K., Bouwknegt, M., Willems, K., de Roda-Husman, A., et al., 2012: Occurrence of Human Enteric Viruses in Commercial Mussels at Retail Level in Three European Countries. *Food and environmental virology,* **4,** 73-80.

Doceul, V., Bagdassarian, E., Demange, A. and Pavio, N., 2016: Zoonotic Hepatitis E Virus: Classification, Animal Reservoirs and Transmission Routes. *Viruses-Basel,* **8**.

Drexler, J.F., Corman, V.M. and Drosten, C., 2014: Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral research,* **101,** 45-56.

Drexler, J.F., Gloza-Rausch, F., Glende, J., Corman, V.M., Muth, D., Goettsche, M., et al., 2010: Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *Journal of Virology,* **84,** 11336-11349.

Drobeniuc, J., Greene-Montfort, T., Le, N.T., Mixson-Hayden, T.R., Ganova-Raeva, L., Dong, C., et al., 2013: Laboratory-based Surveillance for Hepatitis E Virus Infection, United States, 2005-2012. *Emerging infectious diseases,* **19,** 218-222.

Drosten, C., Seilmaier, M., Corman, V.M., Hartmann, W., Scheible, G., Sack, S., et al., 2013: Clinical features and virological analysis of a case of Middle East respiratory syndrome coronavirus infection. *The Lancet. Infectious diseases,* **13,** 745-751.

Feagins, A.R., Opriessnig, T., Guenette, D.K., Halbur, P.G. and Meng, X.J., 2007: Detection and characterization of infectious Hepatitis E virus from commercial pig livers sold in local grocery stores in the USA. *The Journal of General Virology,* **88,** 912-917.

Fehr, A.R. and Perlman, S., 2015: Coronaviruses: an overview of their replication and pathogenesis. *Methods in Molecular Biology,* **1282,** 1-23.

Festa, S., Garbuglia, A.R., Baccini, F., Panzuto, F., Capobianchi, M.R., Santino, I., et al., 2014: Acute fulminant hepatitis E virus genotype 3e infection: Description of the first case in Europe. *Scandinavian journal of infectious diseases,* **46,** 727-731.

Feurer, C., Le Roux, A., Rossel, R., Barnaud, E., Dumarest, M., Garry, P., et al., 2018: High load of hepatitis E viral RNA in pork livers but absence in pork muscle at French slaughterhouses. *International journal of food microbiology,* **264,** 25-30.

Forni, D., Filippi, G., Cagliani, R., De Gioia, L., Pozzoli, U., Al-Daghri, N., et al., 2015: The heptad repeat region is a major selection target in MERS-CoV and related coronaviruses. *Scientific reports,* **5**.

Garbuglia, A.R., Alessandrini, A.I., Pavio, N., Tesse, S., Grignolo, S., Viscoli, C., et al., 2015: Male patient with acute hepatitis E in Genoa, Italy: figatelli (pork liver sausage) as probable source of the infection. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases,* **21,** e4-6.

Geng, Y., Zhao, C., Huang, W., Wang, X., Xu, Y., Wu, D., et al., 2018: Hepatitis E virus was not detected in feces and milk of cows in Hebei province of China: No evidence for HEV prevalence in cows. *International Journal of Food Microbiology,* **291,** 5-9.

Gerna, G., Campanini, G., Rovida, F., Percivalle, E., Sarasini, A., Marchi, A., et al., 2006: Genetic variability of human coronavirus OC43-, 229E-, and NL63-like strains and their association with lower respiratory tract infections of hospitalized infants and immunocompromised patients. *Journal of medical virology,* **78,** 938-949.

Giordani, M.T., Fabris, P., Brunetti, E., Goblirsch, S. and Romano, L., 2013: Hepatitis E and lymphocytic leukemia in Man, Italy. *Emerging infectious diseases,* **19,** 2054-2056.

Gloza-Rausch, F., Ipsen, A., Seebens, A., Gottsche, M., Panning, M., Drexler, J.F., et al., 2008: Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. *Emerging infectious diseases,* **14,** 626-631.

Goel, A. and Aggarwal, R., 2016: Advances in hepatitis E - II: Epidemiology, clinical manifestations, treatment and prevention. *Expert review of gastroenterology & hepatology,* **10,** 1065-1074.

Goffard, A., Demanche, C., Arthur, L., Pincon, C., Michaux, J. and Dubuisson, J., 2015: Alphacoronaviruses Detected in French Bats Are Phylogeographically Linked to Coronaviruses of European Bats. *Viruses-Basel,* **7,** 6279-6290.

Gouilh, M.A., Puechmaille, S.J., Diancourt, L., Vandenbogaert, M., Serra-Cobo, J., Roig, M.L., et al., 2018: SARS-CoV related Betacoronavirus and diverse Alphacoronavirus members found in western old-world. *Virology,* **517,** 88-97.

Graham, R.L. and Baric, R.S., 2010: Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-Species Transmission. *Journal of Virology,* **84,** 3134-3146.

Guan, Y., Peiris, J.S.M., Zheng, B., Poon, L.L.M., Chan, K.H., Zeng, F.Y., et al., 2004: Molecular epidemiology of the novel coronavirus that causes severe acute respiratory syndrome. *Lancet,* **363,** 99-104.

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., et al., 2003: Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. *Science,* **302,** 276-278.

Hara, Y., Terada, Y., Yonemitsu, K., Shimoda, H., Noguchi, K., Suzuki, K., et al., 2014: High prevalence of hepatitis E virus in wild boar (Sus scrofa) in Yamaguchi Prefecture, Japan. *Journal of Wildlife Diseases,* **50,** 378-383.

Harrison, L.C. and DiCaprio, E., 2018: Hepatitis E Virus: An Emerging Foodborne Pathogen. *Frontiers in Sustainable Food Systems,* **2**.

Hartl, J., Otto, B., Madden, R.G., Webb, G., Woolson, K.L., Kriston, L., et al., 2016: Hepatitis E Seroprevalence in Europe: A Meta-Analysis. *Viruses-Basel,* **8**.

Heldt, F.H., Staggmeier, R., Gularte, J.S., Demoliner, M., Henzel, A. and Spilki, F.R., 2016: Hepatitis E Virus in Surface Water, Sediments, and Pork Products Marketed in Southern Brazil. *Food and environmental virology,* **8,** 200-205.

Hemida, M.G., Al-Naeem, A., Perera, R.A., Chin, A.W., Poon, L.L. and Peiris, M., 2015: Lack of middle East respiratory syndrome coronavirus transmission from infected camels. *Emerging infectious diseases,* **21,** 699-701.

Hon, C.C., Lam, T.Y., Shi, Z.L., Drummond, A.J., Yip, C.W., Zeng, F., et al., 2008: Evidence of the recombinant origin of a bat severe acute respiratory syndrome (SARS)-like coronavirus and its implications on the direct ancestor of SARS coronavirus. *Journal of Virology,* **82,** 1819-1826.

Huang, F., Li, Y.L., Yu, W.H., Jing, S.R., Wang, J., Long, F.Y., et al., 2016: Excretion of Infectious Hepatitis E Virus Into Milk in Cows Imposes High Risks of Zoonosis. *Hepatology,* **64,** 350-359.

Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J.C., et al., 2012: Evidence supporting a zoonotic origin of human coronavirus strain NL63. *Journal of Virology,* **86,** 12816-12825.

Ithete, N.L., Stoffberg, S., Corman, V.M., Cottontail, V.M., Richards, L.R., Schoeman, M.C., et al., 2013: Close Relative of Human Middle East Respiratory Syndrome Coronavirus in Bat, South Africa. *Emerging infectious diseases,* **19,** 1697-1699.

Jothikumar, N., Cromeans, T.L., Robertson, B.H., Meng, X.J. and Hill, V.R., 2006: A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. *Journal of virological methods,* **131,** 65-71.

Kamar, N., Bendall, R., Legrand-Abravanel, F., Xia, N.S., Ijaz, S., Izopet, J., et al., 2012: Hepatitis E. *Lancet,* **379,** 2477-2488.

Kamar, N., Izopet, J., Cintas, P., Garrouste, C., Uro-Coste, E., Cointault, O., et al., 2010: Hepatitis E virus-induced neurological symptoms in a kidney-transplant patient with chronic hepatitis. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons,* **10,** 1321-1324.

Kamar, N., Izopet, J., Pavio, N., Aggarwal, R., Labrique, A., Wedemeyer, H., et al., 2017: Hepatitis E virus infection. *Nature Reviews Disease Primers,* **3**.

Kan, B., Wang, M., Jing, H.Q., Xu, H.F., Jiang, X.G., Yan, M.Y., et al., 2005: Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. *Journal of Virology,* **79,** 11892-11900.

Kemenesi, G., Dallos, B., Gorfol, T., Boldogh, S., Estok, P., Kurucz, K., et al., 2014: Molecular Survey of RNA Viruses in Hungarian Bats: Discovering Novel Astroviruses, Coronaviruses, and Caliciviruses. *Vector-Borne and Zoonotic Diseases,* **14,** 846-855.

Khuroo, M.S. and Khuroo, N.S., 2016a: Hepatitis E: Discovery, global impact, control and cure. *World journal of gastroenterology,* **22,** 7030-7045.

Khuroo, M.S. and Khuroo, N.S., 2016b: Transmission of Hepatitis E Virus in Developing Countries. *Viruses,* **8**.

Kilpatrick, A.M. and Randolph, S.E., 2012: Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. *Lancet,* **380,** 1946-1955.

Kim, Y., Cheon, S., Min, C.K., Sohn, K.M., Kang, Y.J., Cha, Y.J., et al., 2016: Spread of Mutant Middle East Respiratory Syndrome Coronavirus with Reduced Affinity to Human CD26 during the South Korean Outbreak. *mBio,* **7**.

Kokkinos, P., Kozyra, I., Lazic, S., Bouwknegt, M., Rutjes, S., Willems, K., et al., 2012: Harmonised Investigation of the Occurrence of Human Enteric Viruses in the Leafy Green Vegetable Supply Chain in Three European Countries. *Food and environmental virology,* **4,** 179-191.

Kukielka, D., Rodriguez-Prieto, V., Vicente, J. and Sanchez-Vizcaino, J.M., 2016: Constant Hepatitis E Virus (HEV) Circulation in Wild Boar and Red Deer in Spain: An Increasing Concern Source of HEV Zoonotic Transmission. *Transboundary and emerging diseases,* **63,** e360-368.

Kumar, S., Subhadra, S., Singh, B. and Panda, B.K., 2013: Hepatitis E virus: the current scenario. *International Journal of Infectious Diseases,* **17,** E228-E233.

Lam, T.T., Hon, C.C., Lam, P.Y., Yip, C.W., Zeng, F. and Leung, F.C., 2008: Comments to the predecessor of human SARS coronavirus in 2003-2004 epidemic. *Veterinary microbiology,* **126,** 390-393.

Lapa, D., Capobianchi, M.R. and Garbuglia, A.R., 2015: Epidemiology of Hepatitis E Virus in European Countries. *International journal of molecular sciences,* **16,** 25711-25743.

Lau, S.K., Woo, P.C., Li, K.S., Huang, Y., Tsoi, H.W., Wong, B.H., et al., 2005: Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proceedings of the National Academy of Sciences of the United States of America,* **102,** 14040-14045.

Lauber, C., Goeman, J.J., Parquet, M.D., Nga, P.T., Snijder, E.J., Morita, K., et al., 2013: The Footprint of Genome Architecture in the Largest Genome Expansion in RNA Viruses. *PLoS pathogens,* **9**.

Lee, G.H., Tan, B.H., Teo, E.C., Lim, S.G., Dan, Y.Y., Wee, A., et al., 2016a: Chronic Infection With Camelid Hepatitis E Virus in a Liver Transplant Recipient Who Regularly Consumes Camel Meat and Milk. *Gastroenterology,* **150,** 355-357 e353.

Lelli, D., Papetti, A., Sabelli, C., Rosti, E., Moreno, A. and Boniotti, M.B., 2013: Detection of Coronaviruses in Bats of Various Species in Italy. *Viruses-Basel,* **5,** 2679-2689.

Leopardi, S., Holmes, E.C., Gastaldelli, M., Tassoni, L., Priori, P., Scaravelli, D., et al., 2018: Interplay between co-divergence and cross-species transmission in the evolutionary history of bat coronaviruses. *Infection Genetics and Evolution,* **58,** 279-289.

Li, S., Liu, M.X., Cong, J.J., Zhou, Y.F. and Miao, Z.M., 2017: Detection and Characterization of Hepatitis E Virus in Goats at Slaughterhouse in Tai'an Region, China. *BioMed Research International*.

Li, T.C., Chijiwa, K., Sera, N., Ishibashi, T., Etoh, Y., Shinohara, Y., et al., 2005a: Hepatitis E virus transmission from wild boar meat. *Emerging infectious diseases,* **11,** 1958-1960.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., et al., 2005b: Bats are natural reservoirs of SARS-like coronaviruses. *Science,* **310,** 676-679.

Lim, Y.X., Ng, Y.L., Tam, J.P. and Liu, D.X., 2016: Human Coronaviruses: A Review of Virus-Host Interactions. *Diseases,* **4**.

Lin, Q., Jiang, J., Li, T., Zhang, Y., Zhang, Z., Zheng, M., et al., 2016: [Expression,Purification,Structure Determination and Immunogenicity Assay of Hepatitis E Virus Capsid Protein p495 Derived from Baculovirus-based Insect Cell]. *Bing du xue bao = Chinese journal of virology,* **32,** 342-348.

Liu, D.X., Fung, T.S., Chong, K.K.L., Shukla, A. and Hilgenfeld, R., 2014: Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral research,* **109,** 97-109.

Long, F.Y., Yu, W.H., Yang, C.C., Wang, J., Li, Y.L., Li, Y., et al., 2017: High prevalence of hepatitis E virus infection in goats. *Journal of Medical Virology,* **89,** 1981-1987.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., et al., 2012: Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet,* **380,** 2095-2128.

Lu, L., Li, C. and Hagedorn, C.H., 2006: Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. *Reviews in Medical Virology,* **16,** 5-36.

Luke, T., Wu, H., Zhao, J., Channappanavar, R., Coleman, C.M., Jiao, J.A., et al., 2016: Human polyclonal immunoglobulin G from transchromosomic bovines inhibits MERS-CoV in vivo. *Science translational medicine,* **8,** 326ra321.

Manrubia, S., Lázaro, E, 2006: Viral evolution, Physics of Life Reviews. **3,** 65-92.

Mansuy, J.M., Peron, J.M., Abravanel, F., Poirson, H., Dubois, M., Miedouge, M., et al., 2004: Hepatitis e in the south west of France in individuals who have never visited an endemic area. *Journal of medical virology,* **74,** 419-424.

Martelli, F., Caprioli, A., Zengarini, M., Marata, A., Fiegna, C., Di Bartolo, I., et al., 2008: Detection of Hepatitis E virus (HEV) in a demographic managed wild boar (Sus scrofa scrofa) population in Italy. *Veterinary microbiology,* **126,** 74-81.

Martinelli, N., Pavoni, E., Filogari, D., Ferrari, N., Chiari, M., Canelli, E., et al., 2015: Hepatitis E virus in wild boar in the central northern part of Italy. *Transboundary and emerging diseases,* **62,** 217-222.

Masters, P.S., 2006: The molecular biology of coronaviruses. *Advances in Virus Research, Vol 66,* **66,** 193-292.

Matsubayashi, K., Kang, J.H., Sakata, H., Takahashi, K., Shindo, M., Kato, M., et al., 2008: A case of transfusion-transmitted hepatitis E caused by blood from a donor infected with hepatitis E virus via zoonotic food-borne route. *Transfusion,* **48,** 1368-1375.

Mazzei, M., Nardini, R., Verin, R., Forzan, M., Poli, A. and Tolari, F., 2015: Serologic and molecular survey for hepatitis E virus in wild boar (Sus scrofa) in Central Italy. *New microbes and new infections,* **7,** 41-47.

Memish, Z.A., Mishra, N., Olival, K.J., Fagbo, S.F., Kapoor, V., Epstein, J.H., et al., 2013: Middle East Respiratory Syndrome Coronavirus in Bats, Saudi Arabia. *Emerging infectious diseases,* **19,** 1819-1823.

Meng, X.J., 2000: Novel strains of hepatitis E virus identified from humans and other animal species: is hepatitis E a zoonosis? *Journal of hepatology,* **33,** 842-845.

Meng, X.J., Purcell, R.H., Halbur, P.G., Lehman, J.R., Webb, D.M., Tsareva, T.S., et al., 1997: A novel virus in swine is closely related to the human hepatitis E virus. *Proceedings of the National Academy of Sciences of the United States of America,* **94,** 9860-9865.

Mesquita, J.R., Oliveira, D., Rivadulla, E., Abreu-Silva, J., Varela, M.F., Romalde, J.L., et al., 2016: Hepatitis E virus genotype 3 in mussels (Mytilus galloprovinciallis), Spain. *Food Microbiology,* **58,** 13-15.

Millet, J.K. and Whittaker, G.R., 2014: Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proceedings of the National Academy of Sciences of the United States of America,* **111,** 15214-15219.

Minosse, C., Selleri, M., Zaniratti, M.S., Cappiello, G., Spano, A., Schifano, E., et al., 2008: Phylogenetic analysis of human coronavirus NL63 circulating in Italy. *Journal of Clinical Virology,* **43,** 114-119.

Monchatre-Leroy, E., Boue, F., Boucher, J.M., Renault, C., Moutou, F., Gouilh, M.A., et al., 2017: Identification of Alpha and Beta Coronavirus in Wildlife Species in France: Bats, Rodents, Rabbits, and Hedgehogs. *Viruses-Basel,* **9**.

Montagnaro, S., De Martinis, C., Sasso, S., Ciarcia, R., Damiano, S., Auletta, L., et al., 2015: Viral and Antibody Prevalence of Hepatitis E in European Wild Boars (Sus scrofa) and Hunters at Zoonotic Risk in the Latium Region. *Journal of comparative pathology,* **153,** 1-8.

Muller, M.A., Meyer, B., Corman, V.M., Al-Masri, M., Turkestani, A., Ritz, D., et al., 2015: Presence of Middle East respiratory syndrome coronavirus antibodies in Saudi Arabia: a nationwide, cross-sectional, serological study. *The Lancet. Infectious diseases,* **15,** 559-564.

Murrison, L.B. and Sherman, K.E., 2017: The Enigma of Hepatitis E Virus. *Gastroenterology & hepatology,* **13,** 484-491.

Mykytczuk, O., Harlow, J., Bidawid, S., Corneau, N. and Nasheri, N., 2017: Prevalence and Molecular Characterization of the Hepatitis E Virus in Retail Pork Products Marketed in Canada. *Food and environmental virology,* **9,** 208-218.

Nagashima, S., Takahashi, M., Kobayashi, T., Nishizawa, T., Nishiyama, T., Primadharsini, P.P., et al., 2017: Characterization of the Quasi-Enveloped Hepatitis E Virus Particles Released by the Cellular Exosomal Pathway. *Journal of Virology,* **91**.

Okamoto, H., Takahashi, M., Nishizawa, T., Fukai, K., Muramatsu, U. and Yoshikawa, A., 2001: Analysis of the complete genome of indigenous swine hepatitis E virus isolated in Japan. *Biochemical and biophysical research communications,* **289,** 929-936.

Omrani, A.S., Al-Tawfiq, J.A. and Memish, Z.A., 2015: Middle East respiratory syndrome coronavirus (MERS-CoV): animal to human interaction. *Pathogens and global health,* **109,** 354-362.

Pauly, M., Pir, J.B., Loesch, C., Sausy, A., Snoeck, C.J., Hubschen, J.M., et al., 2017: Novel Alphacoronaviruses and Paramyxoviruses Cocirculate with Type 1 and Severe Acute Respiratory System (SARS)-Related Betacoronaviruses in Synanthropic Bats of Luxembourg. *Applied and environmental microbiology,* **83**.

Pavio, N., Doceul, V., Bagdassarian, E. and Johne, R., 2017: Recent knowledge on hepatitis E virus in Suidae reservoirs and transmission routes to human. *Veterinary research,* **48**.

Pavio, N., Meng, X.J. and Doceul, V., 2015: Zoonotic origin of hepatitis E. *Current opinion in virology,* **10,** 34-41.

Pavio, N., Merbah, T. and Thebault, A., 2014: Frequent Hepatitis E Virus Contamination in Food Containing Raw Pork Liver, France. *Emerging infectious diseases,* **20,** 1925-1927.

Perlman, S. and Netland, J., 2009: Coronaviruses post-SARS: update on replication and pathogenesis. *Nature Reviews Microbiology,* **7,** 439-450.

Pfefferle, S., Oppong, S., Drexler, J.F., Gloza-Rausch, F., Ipsen, A., Seebens, A., et al., 2009: Distant Relatives of Severe Acute Respiratory Syndrome Coronavirus and Close Relatives of Human Coronavirus 229E in Bats, Ghana. *Emerging infectious diseases,* **15,** 1377-1384.

Qian, Z.H., Dominguez, S.R. and Holmes, K.V., 2013: Role of the Spike Glycoprotein of Human Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in Virus Entry and Syncytia Formation. *PloS one,* **8**.

Renou, C., Afonso, A.M.R. and Pavio, N., 2014: Foodborne Transmission of Hepatitis E Virus from Raw Pork Liver Sausage, France. *Emerging infectious diseases,* **20,** 1945-1947.

Reusken, C.B., Farag, E.A., Jonges, M., Godeke, G.J., El-Sayed, A.M., Pas, S.D., et al., 2014: Middle East respiratory syndrome coronavirus (MERS-CoV) RNA and neutralising antibodies in milk collected according to local customs from dromedary camels, Qatar, April 2014. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin,* **19**.

Reusken, C.B.E.M., Lina, P.H.C., Pielaat, A., de Vries, A., Dam-Deisz, C., Adema, J., et al., 2010: Circulation of Group 2 Coronaviruses in a Bat Species Common to Urban Areas in Western Europe. *Vector-Borne and Zoonotic Diseases,* **10,** 785-791.

Reusken, C.B.E.M., Raj, V.S., Koopmans, M.P. and Haagmans, B.L., 2016: Cross host transmission in the emergence of MERS coronavirus. *Current opinion in virology,* **16,** 55-62.

Ricci, A., Allende, A., Bolton, D., Chemaly, M., Davies, R., Escamez, P.S.F., et al., 2017: Public health risks associated with hepatitis E virus (HEV) as a food-borne pathogen. *EFSA Journal,* **15**.

Rivero-Juarez, A., Frias, M., Martinez-Peinado, A., Risalde, M.A., Rodriguez-Cano, D., Camacho, A., et al., 2017: Familial Hepatitis E Outbreak Linked to Wild Boar Meat Consumption. *Zoonoses and public health,* **64,** 561-565.

Rizzo, F., Edenborough, K.M., Toffoli, R., Culasso, P., Zoppi, S., Dondo, A., et al., 2017: Coronavirus and paramyxovirus in bats from Northwest Italy. *BMC veterinary research,* **13,** 396.

Romano, L., Paladini, S., Tagliacarne, C., Canuti, M., Bianchi, S. and Zanetti, A.R., 2011: Hepatitis E in Italy: A long-term prospective study. *Journal of hepatology,* **54,** 34-40.

Ruggeri, F.M., Di Bartolo, I., Ponterio, E., Angeloni, G., Trevisani, M. and Ostanello, F., 2013: Zoonotic transmission of hepatitis E virus in industrialized countries. *The new microbiologica,* **36,** 331-344.

Said, B., Ijaz, S., Kafatos, G., Booth, L., Thomas, H.L., Walsh, A., et al., 2009: Hepatitis E Outbreak on Cruise Ship. *Emerging infectious diseases,* **15,** 1738-1744.

Sato, Y., Sato, H., Naka, K., Furuya, S., Tsukiji, H., Kitagawa, K., et al., 2011: A nationwide survey of hepatitis E virus (HEV) infection in wild boars in Japan: identification of boar HEV strains of genotypes 3 and 4 and unrecognized genotypes. *Archives of virology,* **156,** 1345-1358.

Schlosser, J., Eiden, M., Vina-Rodriguez, A., Fast, C., Dremsek, P., Lange, E., et al., 2014: Natural and experimental hepatitis E virus genotype 3-infection in European wild boar is transmissible to domestic pigs. *Veterinary research,* **45,** 121.

Serracca, L., Battistini, R., Rossini, I., Mignone, W., Peletto, S., Boin, C., et al., 2015: Molecular Investigation on the Presence of Hepatitis E Virus (HEV) in Wild Game in North-Western Italy. *Food and environmental virology,* **7,** 206-212.

Sethna, P.B., Hofmann, M.A. and Brian, D.A., 1991: Minus-strand copies of replicating coronavirus mRNAs contain antileaders. *Journal of Virology,* **65,** 320-325.

Simmons, G., Gosalia, D.N., Rennekamp, A.J., Reeves, J.D., Diamond, S.L. and Bates, P., 2005: Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proceedings of the National Academy of Sciences of the United States of America,* **102,** 11876-11881.

Smith, D.B., Simmonds, P., Izopet, J., Oliveira-Filho, E.F., Ulrich, R.G., Johne, R., et al., 2016: Proposed reference sequences for hepatitis E virus subtypes. *The Journal of general virology,* **97,** 537-542.

Smith, D.B., Simmonds, P., Jameel, S., Emerson, S.U., Harrison, T.J., Meng, X.J., et al., 2014: Consensus proposals for classification of the family Hepeviridae. *The Journal of General Virology,* **95,** 2223-2232.

Smith, J.L., 2001: A review of hepatitis E virus. *Journal of food protection,* **64,** 572-586.

Song, H.D., Tu, C.C., Zhang, G.W., Wang, S.Y., Zheng, K., Lei, L.C., et al., 2005: Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proceedings of the National Academy of Sciences of the United States of America,* **102,** 2430-2435.

Sonoda, H., Abe, M., Sugimoto, T., Sato, Y., Bando, M., Fukui, E., et al., 2004: Prevalence of hepatitis E virus (HEV) Infection in wild boars and deer and genetic identification of a genotype 3 HEV from a boar in Japan. *Journal of clinical microbiology,* **42,** 5371-5374.

Spada, E., Pupella, S., Pisani, G., Bruni, R., Chionne, P., Madonna, E., et al., 2018: A nationwide retrospective study on prevalence of hepatitis E virus infection in Italian blood donors. *Blood Transfusion-Italy,* **16,** 413-421.

Spahr, C., Knauf-Witzens, T., Vahlenkamp, T., Ulrich, R.G. and Johne, R., 2018: Hepatitis E virus and related viruses in wild, domestic and zoo animals: A review. *Zoonoses and Public Health,* **65,** 11-29.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C.K., Zhou, J., et al., 2016: Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends in Microbiology,* **24,** 490-502.

Takahashi, K., Kitajima, N., Abe, N. and Mishiro, S., 2004: Complete or near-complete nucleotide sequences of hepatitis E virus genome recovered from a wild boar, a deer, and four patients who ate the deer. *Virology,* **330,** 501-505.

Takahashi, M., Nishizawa, T., Nagashima, S., Jirintai, S., Kawakami, M., Sonoda, Y., et al., 2014: Molecular characterization of a novel hepatitis E virus (HEV) strain obtained from a wild boar in Japan that is highly divergent from the previously recognized HEV strains. *Virus research,* **180,** 59-69.

Tao, Y., Tang, K., Shi, M., Conrardy, C., Li, K.S., Lau, S.K., et al., 2012: Genomic characterization of seven distinct bat coronaviruses in Kenya. *Virus research,* **167,** 67-73.

Taylor, L.H., Latham, S.M. and Woolhouse, M.E., 2001: Risk factors for human disease emergence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences,* **356,** 983-989.

Tei, S., Kitajima, N., Ohara, S., Inoue, Y., Miki, M., Yamatani, T., et al., 2004: Consumption of uncooked deer meat as a risk factor for hepatitis E virus infection: An age- and sex-matched case-control study. *Journal of Medical Virology,* **74,** 67-70.

Tei, S., Kitajima, N., Takahashi, K. and Mishiro, S., 2003: Zoonotic transmission of hepatitis E virus from deer to human beings. *Lancet,* **362,** 371-373.

Teo, C.G., 2010: Much meat, much malady: changing perceptions of the epidemiology of hepatitis E. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases,* **16,** 24-32.

Thiry, D., Mauroy, A., Pavio, N., Purdy, M.A., Rose, N., Thiry, E., et al., 2017: Hepatitis E Virus and Related Viruses in Animals. *Transboundary and emerging diseases,* **64,** 37-52.

van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T.M., Raj, V.S., Zaki, A.M., et al., 2012: Genomic Characterization of a Newly Discovered Coronavirus Associated with Acute Respiratory Distress Syndrome in Humans. *mBio,* **3**.

Vijayanand, P., Wilkins, E. and Woodhead, M., 2004: Severe acute respiratory syndrome (SARS): a review. *Clinical Medicine,* **4,** 152-160.

Vijaykrishna, D., Smith, G.J.D., Zhang, J.X., Peiris, J.S.M., Chen, H. and Guan, Y., 2007: Evolutionary insights into the ecology of coronaviruses. *Journal of Virology,* **81,** 4012-4020.

Vina-Rodriguez, A., Schlosser, J., Becher, D., Kaden, V., Groschup, M.H. and Eiden, M., 2015: Hepatitis E virus genotype 3 diversity: phylogenetic analysis and presence of subtype 3b in wild boar in Europe. *Viruses,* **7,** 2704-2726.

Walachowski, S., Dorenlor, V., Lefevre, J., Lunazzi, A., Eono, F., Merbah, T., et al., 2014: Risk factors associated with the presence of hepatitis E virus in livers and seroprevalence in slaughter-age pigs: a retrospective study of 90 swine farms in France. *Epidemiology and infection,* **142,** 1934-1944.

Wang, M., Yan, M., Xu, H., Liang, W., Kan, B., Zheng, B., et al., 2005a: SARS-CoV infection in a restaurant from palm civet. *Emerging infectious diseases,* **11,** 1860-1865.

Wang, Q., Qi, J., Yuan, Y., Xuan, Y., Han, P., Wan, Y., et al., 2014: Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. *Cell host & microbe,* **16,** 328-337.

Wang, Z.G., Zheng, Z.H., Shang, L., Li, L.J., Cong, L.M., Feng, M.G., et al., 2005b: Molecular evolution and multilocus sequence typing of 145 strains of SARS-CoV. *Febs Letters,* **579,** 4928-4936.

Wenzel, J.J., Preiss, J., Schemmerer, M., Huber, B., Plentz, A. and Jilg, W., 2011: Detection of hepatitis E virus (HEV) from porcine livers in Southeastern Germany and high sequence homology to human HEV isolates. *Journal of Clinical Virology,* **52,** 50-54.

Wernery, U., 2014: Some Epidemiological Studies on Mers Coronavirus in Dromedaries in the United Arab Emirates- a Short Communication. *Journal of Camel Practice and Research,* **21,** 1-4.

Woo, P.C.Y., Lau, S.K.P., Teng, J.L.L., Cao, K.Y., Wernery, U., Schountz, T., et al., 2016: New Hepatitis E Virus Genotype in Bactrian Camels, Xinjiang, China, 2013. *Emerging infectious diseases,* **22,** 2219-2221.

Woolhouse, M.E.J. and Gowtage-Sequeria, S., 2005: Host range and emerging and reemerging pathogens. *Emerging infectious diseases,* **11,** 1842-1847.

Wu, D.L., Tu, C.C., Xin, C., Xuan, H., Meng, Q.W., Liu, Y.G., et al., 2005: Civets are equally susceptible to experimental infection by two different severe acute respiratory syndrome coronavirus isolates. *Journal of Virology,* **79,** 2620-2625.

Wu, J., Si, F., Jiang, C., Li, T. and Jin, M., 2015: Molecular detection of hepatitis E virus in sheep from southern Xinjiang, China. *Virus genes,* **50,** 410-417.

Yan, B.Y., Zhang, L., Gong, L.F., Lv, J.J., Feng, Y., Liu, J.Y., et al., 2016: Hepatitis E Virus in Yellow Cattle, Shandong, Eastern China. *Emerging Infectious Diseases,* **22,** 2211-2212.

Yang, L., Wu, Z.Q., Ren, X.W., Yang, F., Zhang, J.P., He, G.M., et al., 2014: MERS-Related Betacoronavirus in Vespertilio superans Bats, China. *Emerging infectious diseases,* **20,** 1260-1262.

Yazaki, Y., Mizuo, H., Takahashi, M., Nishizawa, T., Sasaki, N., Gotanda, Y., et al., 2003: Sporadic acute or fulminant hepatitis E in Hokkaido, Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. *The Journal of General Virology,* **84,** 2351-2357.

Yin, X., Ambardekar, C., Lu, Y. and Feng, Z., 2016: Distinct Entry Mechanisms for Nonenveloped and Quasi-Enveloped Hepatitis E Viruses. *Journal of Virology,* **90,** 4232-4242.

Ying, T., Du, L., Ju, T.W., Prabakaran, P., Lau, C.C., Lu, L., et al., 2014: Exceptionally potent neutralization of Middle East respiratory syndrome coronavirus by human monoclonal antibodies. *Journal of Virology,* **88,** 7796-7805.

Yip, C.W., Hon, C.C., Shi, M., Lam, T.T., Chow, K.Y., Zeng, F., et al., 2009: Phylogenetic perspectives on the epidemiology and origins of SARS and SARS-like coronaviruses. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases,* **9,** 1185-1196.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D. and Fouchier, R.A., 2012: Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *The New England journal of medicine,* **367,** 1814-1820.

Zanetti, A.R., Schlauder, G.G., Romano, L., Tanzi, E., Fabris, P., Dawson, G.J., et al., 1999: Identification of a novel variant of hepatitis E virus in Italy. *Journal of medical virology,* **57,** 356-360.

Zumla, A., Chan, J.F., Azhar, E.I., Hui, D.S. and Yuen, K.Y., 2016: Coronaviruses - drug discovery and therapeutic options. *Nature reviews. Drug discovery,* **15,** 327-347.

CHAPTER 2

Aims of the work

2. Aims of the work

The zoonoses are infection diseases that can be naturally transmitted from animals to humans. Hepatitis E is an emerging disease caused by HEV and is considered a major public health concern worldwide. The epidemiology and the transmission mode of the diseases are different depending on the geographical area. In low income countries, it is a waterborne disease, while in developed countries the virus is transmitted zoonotically from animal reservoirs to humans. Since the last decade, in Europe, an increasing number of human infections, caused by the zoonotic genotypes HEV-3 and less frequently by HEV-4, has been reported.

HEV-3 and HEV-4 infect several animal species asymptomatically such as domestic pigs, considered the main reservoir, wild boars and deer. HEV-3 and HEV-4 are mainly transmitted to humans by consumption of raw or undercooked food of animal origin.

Two other zoonotic viruses SARS-CoV and MERS-CoV, both from the *Coronaviridae* family, emerged in the last decade, causing thousands of human cases of severe respiratory disease worldwide. Phylogenetic analysis on human and animal strains confirmed the cycle of viral transmission from a mammal amplifier host to human and identified the origin of both strains in bats, considered the main reservoir.

HEV and CoV cannot be grown efficiently *in vitro*, and up to date most of the information available, their classification and the data on virus circulation, are based on molecular detection methods (Real-time RT-PCR; PCR) and phylogenetic analyses on viral sequences. New viral strains are continuously identified but some strains are still unclassified. To overcome these limits, the analyses of whole genome sequences are needed. The Next Generation Sequencing (NGS) technique, providing a massively parallel sequencing, allows obtaining a great amount of sequence information and assembling of several complete genomes using a single sequence run.

The aim of the present work was to investigate the presence of HEV strains in animal reservoirs. Furthermore, for both HEV and CoV the evolutionary correlations among strains and their origin by phylogenetic analysis were established.

The first approach was to identify and to sequence the virus strains by both traditional molecular methods and NGS. In Chapter 3, we applied the NGS method to sequence the full genome of two HEV strains detected in two pig fecal samples. The obtained full genomes were used to perform an accurate phylogenetic analysis and leaded to the identification of a novel clade (defined as subtype) of HEV-3, named HEV-3l. In Chapter 4, we described the results of the molecular surveillance on HEV-3 circulation in a wild boar population hunted in Southern Italy. HEV strains, detected from wild boars, were partially classified by sequencing and phylogenetic analyses of short genomic regions. Results showed heterogeneity among strains, including the detection of HEV-3 strains

never detected before in wild boar. Two HEV-3 strains, detected in liver wild boars, were fully sequenced by NGS and subjected to phylogenetic analyses, identifying two novel and uncommon variants of HEV-3 (Chapter 5). The results obtained showed the high heterogeneity of HEV-3 strains in wild boar, small animal populations hosted several strains some of which never detected before in Italy. To assess and better understand the role of this reservoir in the HEV transmission we conducted (Chapter 6) a molecular surveillance on the wild boar population hunted in five small areas from Central Italy (Lazio). Results confirmed the widespread of HEV-3 in wild boars; animals from all but one hunting areas surveilled were positive for HEV.

To better understand the CoV species circulating in Italian bat population, five bats positive for CoV infection, identified in North Italy in a previous surveillance study, were fully sequenced and characterized. In Chapter 7, two CoV complete genome sequences showed correlation with Beta-CoV genera and classified into MERS-CoV species confirming that bats are reservoir for different CoV species. In Chapter 8, three complete genomes of Alpha-CoV genera were classified into two novel CoV species able to infect the same bat species.

For diagnostic and surveillance purpose, the NGS method was implemented using multiplexed PCRs able to amplify HEV-3 and MERS-CoV strains. Results were promising since full genome sequences with high coverage rate were obtained from 4 different subtypes of HEV-3 and from two MERS-CoV like strains. The new method increased the number of specific viral sequences obtaining complete genomes.

CHAPTER 3

Proposal for a new subtype of the zoonotic genotype 3 Hepatitis E virus: HEV-3l

3. Proposal for a new subtype of the zoonotic genotype 3 Hepatitis E virus: HEV-3l

Hepatitis E virus is a pathogen that causes in human acute hepatitis, generally self-limited, that can become chronic in immunocompromised patients. In industrialized countries, most of the cases are linked to the zoonotic HEV-3 and HEV-4 genotypes that infect both humans and animals among which domestic swine and wild boar are the main reservoirs. In Europe, foodborne is the main route of transmission, linked to consumption of raw or undercooked meat or meat products containing liver from pork and wild boar infected by HEV-3 and rarely by HEV-4. In *vitro* cultivation of HEV is still poorly efficient and both detection and classification of the virus are based on molecular methods followed by sequence comparisons and phylogenetic analyses. The wide heterogeneity of HEV determines its classification in genotypes that are further divided in subtypes, named by letter. Molecular methods, such as end-point reverse transcription PCR (RT-PCR), are based on amplification of short genomic fragments conserved among HEV strains that sometimes make virus characterization difficult. The analysis of short genomic sequences is not always sufficient to have a robust classification or characterization of the analyzed strains. A better phylogenetic analysis can be obtained using long or complete genome sequences. With the aim to define rules for HEV classification, some criteria have been recently proposed together with a set of reference genomes belonging to each genotype and further divided in subtypes, to be used for comparisons. The classification by complete genome sequences is based on the phylogenetic clustering with the HEV reference strains. However, a cut-off nucleotide distance value has not been established for HEV-3 yet. Two HEV-3 strains detected from pig feces, were not classifiable in any of the subtypes define so far by sequence comparison using the conserved short genome regions. To obtain a longer genome sequences to be used for classification, the two samples were subjected to NGS method developed in this study, based on random amplification of all nucleic acids followed by sequencing. The method failed in HEV complete genome amplification and the sequences obtained by NGS were used to draw specific primer pairs used to complete the genome sequences. The two full genome sequences were analyzed by phylogenetic methods and compared to HEV-3 references. Following the HEV classification criteria, the two Italian strains together with a French strain, whose full genome sequence was available online, were classified into a novel HEV-3 subtype, named l, distant from the other described HEV strains. The strains belonging to this new subtype have been identified to date only in Italy and France and could therefore represent a viral subtype that has not been definitively adapted to pigs, circulating moderately in the swine population.

Virus Research 248 (2018) 1–4

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Short communication

Proposal for a new subtype of the zoonotic genotype 3 Hepatitis E virus: HEV-3l

Luca De Sabato[a,b], Philippe Lemey[c,*], Bram Vrancken[c], Lebana Bonfanti[d], Letizia Ceglie[d], Gabriele Vaccari[a,*], Ilaria Di Bartolo[a]

[a] Istituto Superiore di Sanità, Department of Food Safety, Nutrition and Veterinary Public Health, Viale Regina Elena, 299, 00161, Rome, Italy
[b] University Roma Tre, Department of Sciences, Viale Guglielmo Marconi 446, 00146 Rome, Italy
[c] KU Leuven, University of Leuven, Rega Institute, Department of Microbiology and Immunology, Herestraat 49, 3000 Leuven, Belgium
[d] Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università, 10, 35020 Legnaro PD, Italy

ARTICLE INFO

Keywords:
Hepatitis E virus
Complete genome
Next generation sequencing
Genotype 3
Subtype
Swine

ABSTRACT

The near-complete genomic sequences of two hepatitis E virus (HEV) strains, detected from feces of infected pigs, were obtained. Phylogenetic analysis and p-distance comparisons of the complete coding regions showed a close relationship to the French swine strain FR-SHEV3c-like detected in 2006 (*p*-distance value 0.101), belonging to HEV-3 but not assigned to any known subtype. The three HEV sequences showed, relatively high nucleotide distances (*p*-distance > 0.129) compared to the other defined HEV subtype references and unclassified strains. The HEV classification criteria and the high sequence similarity suggest that these strains can be assigned to a putative novel subtype of genotype 3, HEV-3l.

Hepatitis E virus (HEV) causes acute hepatitis, mostly self-limiting, with a mortality rate of 1% that can increase up to 25% in pregnant women. HEV infection can develop into a chronic disease in immunocompromised individuals (Donnelly et al., 2017). HEV is a quasi-enveloped virus with a positive-stranded RNA (Okamoto, 2011). The genome is approximately 7.2 kb and divided in three Open Reading Frames (ORFs): ORF1 encodes the non-structural proteins, ORF2 the capsid protein and ORF3 a small multifunctional phosphoprotein. HEV is classified in the family of *Hepeviridae*, divided into two genera: *Piscihepevirus* that infects trout and *Orthohepevirus* infecting several mammalian and avian species (Smith et al., 2014). The strains that infect humans are included in the *Orthohepevirus A* species and classified into 7 genotypes (HEV1-7) (Smith et al., 2014; Lee et al., 2016; Woo et al., 2016). Genotypes HEV-1 and HEV-2, transmitted by the fecal-oral route, infect humans in developing countries where the disease occurs mainly as outbreaks. HEV-1 and HEV-2 also cause sporadic cases linked to traveling to endemic areas. Genotypes HEV-3 and HEV-4 are foodborne and zoonotic, infecting both humans and animals, and circulate in developed countries causing sporadic cases and small outbreaks (Ricci et al., 2017). The foodborne transmission of the zoonotic HEV, caused by the consumption of raw or undercooked meat and organs, was confirmed by detection of the same viral strains in humans and leftover food derived from swine, wild boar and deer (Doceul et al., 2016). In addition to these more common genotypes, novel reservoirs and genotypes have been described in recent years in rabbit (HEV-3), yak (HEV-4), wild boar (HEV-5, -6) and camel (HEV-7) (Smith et al., 2014; Lee et al., 2016; Woo et al., 2016). The classification of HEV in genotypes is based on a p-distance threshold value of 0.088 for amino acid sequences of the concatenated ORF1, excluding the Hypervariable region (HVR), and ORF2 (Smith et al., 2014). The first classification into 4 genotypes and 24 subtypes was proposed by Lu et al. (2006) and based on the alignment of short genome sequences. The results of more recent classification of *Orthohepevirus A* led to the definition of 7 genotypes including 32 subtypes. Since the cut-off p-distance value for subtype assignment has not been defined yet, the establishment of a novel subtype is based on the clustering in the phylogenetic analysis and the relative high nucleotide identity (low p-distance values) of the full genomes among at least three epidemiologically unrelated strains which showed a high distance with other defined subtypes. Some strains do not fit the subtype definitions and were not assigned (Smith et al., 2016; Miura et al., 2017). Smith et al. (2016) updated previous HEV classification by providing a list of subtype reference strain complete genomes. However, the extensive genomic diversity of HEV strains and the small number of available complete genome ORFs sequences hamper designing a generally applicable subtyping strategy.

In this study, we obtained complete ORFs genome sequences of two swine HEV-3 strains (named SWHEV75BO2012 and HEV/13RS985-5) previously detected in two farms in Northern Italy by Monini et al.

* Corresponding authors.
E-mail addresses: philippe.lemey@kuleuven.be (P. Lemey), gabriele.vaccari@iss.it (G. Vaccari).

https://doi.org/10.1016/j.virusres.2018.02.010
Received 23 January 2018; Received in revised form 9 February 2018; Accepted 12 February 2018
Available online 13 February 2018
0168-1702/

Table 1
Primer sets used for RT-PCRs of the Italian swine HEV strains.

Genomic region	Primer name	Sequence 5′-3′	Position[a]	Product size (bp)	Reference
ORF1	pre-ATG ORF1Fw	ccacrtatgtggtcgatgcc	6–26	1055	Designed in this study
	HEV 1050 Rw	acggagataggtcataagc	1042–1061		Designed in this study
	SwHEV853Fw	gccttatgtcccrtaccc	853–871	1133	Designed in this study
	SwHEV2000Rw	tataaagggcgctgcaaaag	1966–1986		Designed in this study
	SwHEV1354Fw	ccttgtttttgatgaggcggtcc	1370–1393	1509	Di Bartolo et al. (2016)
	SwHEV2893 Rw	gtycgygcdgtytcrttcag	2859–2879		Designed in this study
	SwHEV2744 Fw	cctgggagcgtaaccatcg	2757–2776	991	Di Bartolo et al. (2016)
	Rw3728-orf1HEV	gcagggtagcaagattcctgtc	3725–3748		Designed in this study
ORF1/2/3	Fw3480-orf1HEV	gatgccagggcccttatccaatcg	3527–3551	1829	Designed in this study
	HEVORF2/3con-a1	aggggttggttggatgaatataggg	5331–5356		Erker et al. (1999)
	HEVORF2/3con-s1	gtatcggkykgaatgaataacatgt	5125–5150	867	Erker et al. (1999)
	Rw5960-orf1HEV	cagaggtagcctcttcttcgg	5971–5992		Designed in this study
ORF2	3156	aattatgcycagtaycggrgttg	5712–5735	808	Meng et al. (1997)
	ConsORF2-a1	cttgttcrtgytggttrtcataatc	6495–6520		Wang et al. (1999)
	ConsORF2-s1	gacagaattratttcgtcggctgg	6323–6347	861	Wang et al. (1999)
	HEV7160Rw	aactatgaaggggcacaag	7164–7184		Designed in this study

[a] Referred to strain accession number JQ953664.

Table 2
Percent identity values of the SWHEV75BO2012 and HEV/13RS985-5 strains with HEV-3 genotype reference sequences for HEV Subtypes.

HEV subtype	Accession number	Strain name	Country	SWHEV75BO2012	HEV/13RS985-5
3a	AF082843	Meng	USA	84.5	84.5
3b	AP003430	JRA1	Japan	84	84.1
3c	FJ705359	wbGER27	Germany	85.2	85.2
3d*	AF296165	TW12SW	Taiwan	82.6	82.9
3e	AB248521	swJ8-5	Japan	81.8	81.9
3f	AB369687	E116-YKH98C	Japan	82.1	82.2
3g	AF455784	Osh205	Kyrgyzstan	82.7	82.7
3h	JQ013794	TR19	France	86.3	86.2
3i	FJ998008	BB02	Germany	85.1	85.2
3j	AY115488	Arkell	Canada	83.9	83.6
3k**	–	–	Japan	84.2–84.6	84.2–84.6
3	AB290312	swMN06-A1288	Mongolia	86.6	86.7
3	JQ953664	FR-SHEV3c-like	France	89.9	89.9
3	AB290313	swMN06-C1056	Mongolia	82.3	82.2
3	EU360977	swX07-E1	Sweden	82.2	82.2
3	KJ873911	FR_R	Germany	81.5	81.4
3	KY780957	SW/16-0282	Switzerland	85.5	85.6
3	KU513561	IC2011	Spain	85.1	85
3	KP294371	MWP_2010	Germany	85	85

* Only 304 nt in ORF2 has been reported .

** HEV-3k four strains are available, accession numbers LC176493, LC176492, AB369689, AB740232 (strain names: HE-JA16-057, HE-JA16-0578, E088-STM04C and G3-HEV83-2-27, respectively).

(2015) and by Monne et al. (2013), respectively. Total RNA was extracted from 1 ml of 10% fecal suspension by RNeasy Mini Kit (Qiagen, Milan, Italy) and quantified by RT-qPCR (Real-Time Quantitative Reverse Transcription PCR), as previously described (Di Bartolo et al., 2015). The RNA was analyzed by nested RT-PCR, amplifying two regions in the ORF1: Methyltransferase (MetT: ConsORF2-s1/ConsORF2-a1; ConsORF2-s2/ConsORF2-a2) (Wang et al., 1999), and RNA dependent RNA polymerase (RdRp: EAP pool/ISP pool; IAP pool/ISP pool) (Roth et al., 2016). For both strains short ORF2 sequences were already available (GenBank Accession numbers: KF888274 and KF939866) (Monini et al., 2015; Monne et al., 2013). Amplicons were sequenced with the dideoxy chain-termination method. The RNA was also subjected to Next Generation Sequencing (NGS). The double strand cDNA was prepared following the Sequence Independent Single Primer Amplification (SISPA) method (Djikeng et al., 2008) and the resulting amplicons fragmented to obtain the DNA library (280–320 bp) by Ion Xpres Plus gDNA Fragment Library Preparation (Thermo Fisher Scientific, Rodano, Italy). The libraries were sequenced by the Ion PGM platform (Thermo Fisher Scientific) using Ion PGM sequencing 200 kits v2 (Thermo Fisher Scientific) and loaded on Ion 316 Chip v2 (Thermo Fisher Scientific). The bioinformatic analysis was performed using the online tool Galaxy Aries (https://aries.iss.it/). Reads were cleaned by removing low-quality bases, primer and adaptor sequences. The contigs obtained by *de novo* assembly were used to identify the closest reference genome by BLASTn. The resulting HEV-3 reference strain, FR-SHEV3c-like (JQ953664), was used to assemble the reads by mapping. This led to 4 contigs for SWHEV75BO2012 that range between 350 and 940 bases and cover a total of 3100 nt of ORF1 and ORF2. In contrast, the application of NGS to HEV/13RS985-5 strain did not succeed, resulting in only 8 reads that were not used in the following analysis. Results obtained in this study by NGS were not entirely satisfactory, despites using sample containing 2.8×10^5 (SWHEV75BO2012) or 7.1×10^5 (HEV/13RS985-5) HEV genome equivalents (GE)

The HEV coding regions were obtained by primer walking using 8 sets of primers (Table 1).

Primers were positioned to generate overlapping amplicons spanning the genome regions not sequenced by NGS. The entire coding regions sequence of HEV/13RS985-5 strain was obtained by primer walking (Table 1).

Nucleotide sequences were analyzed edited and assembled using the Bionumerics software V. 6.5 (Applied Maths, Kortrijk, Belgium) and deposited in GenBank NCBI under the accession numbers: KY766999

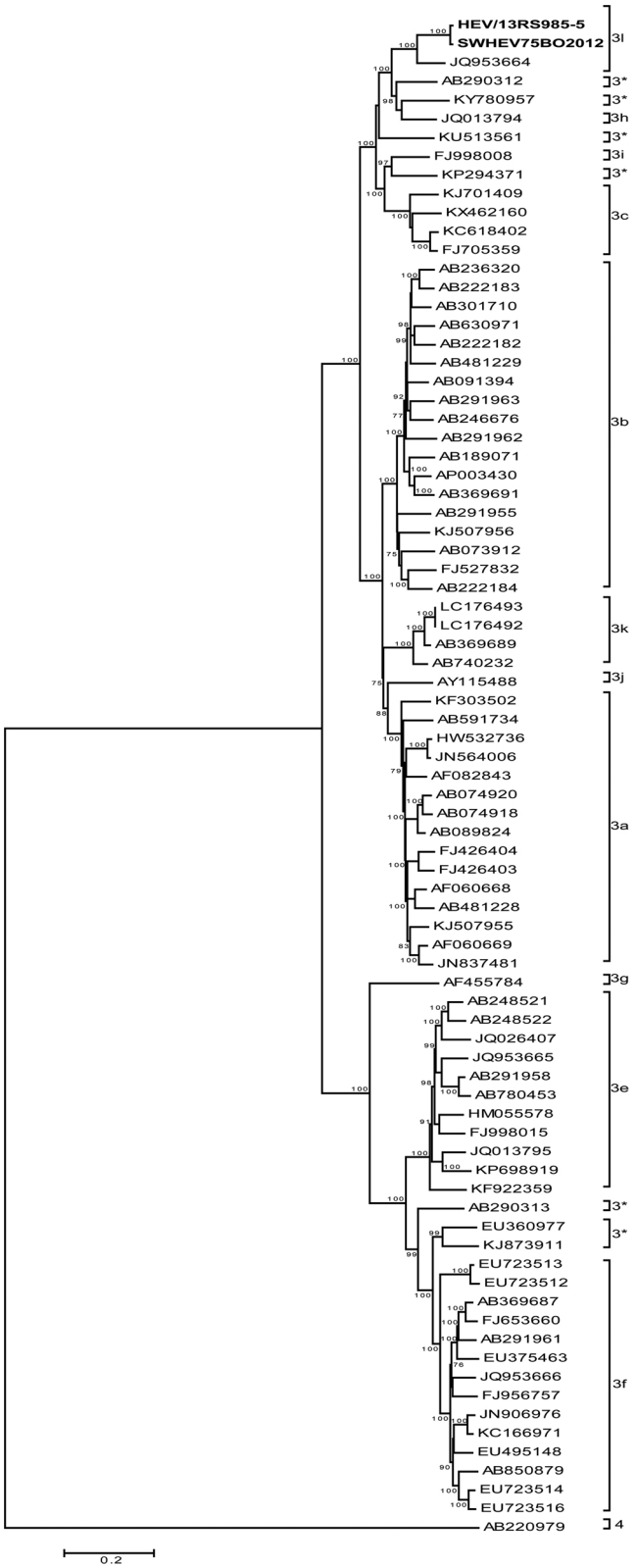


Fig. 1. Maximum likelihood phylogenetic tree reconstruction based on the full-length sequences of 80 HEV-3 strains including the Italian strains SWHEV75BO2012 and HEV/13RS985-5 indicated in bold. The tree was inferred under the GTR + G + I substitution model and a bootstrap resampling process (1000 replications) was used to assess node support. Bootstraps values > 70 are indicated at their respective nodes. Sequences from animals and human strains, belonging to HEV3-a-c and e-k subtypes of genotype 3, have been included in the tree. The HEV-4 strain was used as outgroup.

and MG674164.

To detect the occurrence of recombination, a dataset including full-length sequences of 80 HEV-3 was analyzed by SplitsTree4 (Huson and Bryant, 2006) and RDP4 (Martin et al., 2015) software. The recombination events were assessed using the Phi test of SplitsTree4 and six different methods (GENECONV, BootScan, MaxChi, Chimaera, 3Seq, SiScan) implemented in RDP4 software, using the default settings.

Phylogenetic inference (using maximum likelihood) and calculation of *p*-distance values were performed using MEGA7 (Kumar et al., 2016), and the substitution model suggested by the model-test procedure implemented in MEGA7. Support values for the nodes were obtained by performing 1000 bootstrap analyses.

The two Italian lineages were first tentatively characterized based on sequence similarity (BLASTn) of three short sequence fragments that are commonly amplified for HEV molecular detection (250 nt RdRp, 272 nt MetT and 408 nt of the capsid region) with known strains. SWHEV75BO2012 and HEV/13RS985-5 strains shared > 98% nucleotide identity (nt. id.) with each other and 88–91% nt. id with the French swine strain (FR-SHEV3c-like, JQ953664) in the three short genome regions. It was not possible to confidently classify the Italian strains by phylogenetic analyses due to unresolved and statistically poorly supported nodes in the trees estimated from each of the three short regions (data not shown). For this reason, we obtained and analyzed the complete ORFs genomes sequences of SWHEV75BO2012 and HEV/13RS985-5 (see above). The ORFs sequences were 7130 nt long, organized in a 5112 nt-long (1–5112, 1703 aa) ORF1, a 1983 nt (5147–7129, 660 aa) ORF2, a 369 nt (5109–5477, 123 aa) ORF3. There were 7 unresolved nucleotides in the NGS-based consensus sequence of SWHEV75BO2012. This within-host variability was confirmed by dideoxy chain-termination sequencing, and corresponds to 4 silent and 1 non-synonymous mutations in ORF1 and to 1 silent and 1 non-synonymous change in ORF2. The ORFs genomes sequences of HEV/13RS985-5 presented 4 unresolved sites, corresponding to 2 silent mutations in ORF1 and 2 in ORF2. The Phi test for recombination did not detect statistically significant evidence of recombination ($p = 0.9886$). Results obtained by RDP4 (P values > 0.05) also supported the absence of recombination. The sequences of the two Italian strains showed 99% nt.id. in the three ORFs. The degree of similarity between the two strains was also confirmed at amino acid level (99.2% in the ORF1 and ORF3, 100% in the ORF2). The identity at the amino acidic level was higher than nucleotide, due to several nucleotide changes observed in the third position of codons that confer silent mutations.

The sequences of the Italian strains were compared with reference subtypes sequences proposed by Smith et al., 2016. Results showed that the Italian strains displayed the highest nt. id. 89.9–90% with FR-SHEV3c-like strain and 81.5–86.7% nt. id. with reference sequences assigned to HEV3 subtypes (HEV3a to HEV3k) and some not classified strains (Table 2) (Smith et al., 2016; Miura et al., 2017).

To conclusively identify the subtype of SWHEV75BO2012 and HEV/13RS985-5, two phylogenetic trees were built using reference sequences of HEV-3 subtypes −a to −c and −e to- k (Smith et al., 2016; Miura et al., 2017), excluding (data not shown) or including (Fig. 1) the hypervariable region (HVR) in ORF1 (amino acid residues 706–778 referred to Acc. n°. M73218).The two trees showed the same topology. The phylogeny (Fig. 1) has the expected topology with well-supported branching events separating the HEV-3 genotype strains. The two Italian strains cluster together with the FR-SHEV3c-like strain (JQ953664) in a highly supported clade that is most closely related to the reference HEV–3 h (TR19, JQ013794) and two not classified strains (SW/16-0282, KY780957; swMN06-A1288, AB290312). The calculated nucleotide *p*-distance values among all HEV strains was < 0.120 within subtypes and > 0.118 between subtypes (Fig. 1). The Italian strains showed a *p*-distance value of < 0.101 with FR-SHEV3c-like strain and > 0.129 with the major clade represented by −3c, −3i, −3 h prototype strains and some not classified strains. These results

suggested that the Italian strains and FR-SHEV3c-like strain cluster into the same HEV-3 subtype. Following the latest criterion for HEV classification (Smith et al., 2016; Miura et al., 2017), the high nt. id. among the three strains (two Italian and one French) that are epidemiologically unrelated, their divergence from other subtype reference strains as described above, allow for the confident identification of a novel subtype, named HEV-3l.

The recent work by Smith et al. (2016) provides a set of subtype reference genomes that help to classify new strains with shared rules. This, combined with the analysis of *p*-distances, enables a well-supported classification of the Italian and French swine HEV strains into the novel HEV–3 l subtype. Based on sequence data available on NCBI, the HEV–3 l subtype is very rare, and the reasons for the low prevalence and the limited geographical distribution (Italy and France) of these novel HEV–3 l strains remain unknown. Recently, cell culture system for *in vitro* cultivation has been developed (Okamoto, 2011). However, given the extent heterogeneity of HEV strains, sequencing and phylogenetic analyses represent the main approach to establish, and monitor circulation, host, and geographical distribution of the HEV-3 subtypes.

Acknowledgements

This work was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration under Grant Agreement no. 278433 PREDEMICS. Bram Vrancken is a postdoctoral research fellow funded by the Fonds Wetenschappelijk Onderzoek (FWO, Flanders, Belgium).

References

Di Bartolo, I., Angeloni, G., Ponterio, E., et al., 2015. Detection of hepatitis E virus in pork liver sausages. Int. J. Food Microbiol. 193, 29–33.
Di Bartolo, I., Angeloni, G., Monini, M., et al., 2016. Complete genome sequencing of a genotype 3 hepatitis E virus strain identified in a swine farm in Italy. Virus Res. 211, 89–95.
Djikeng, A., Halpin, R., Kuzmickas, R., et al., 2008. Viral genome sequencing by random priming methods. BMC Genomics 9 (5).
Doceul, V., Bagdassarian, E., Demange, A., et al., 2016. Zoonotic hepatitis E virus: classification, animal reservoirs and transmission routes. Viruses 8 (10).
Donnelly, M.C., Scobie, L., Crossan, C.L., et al., 2017. Review article: hepatitis E-a concise review of virology, epidemiology, clinical presentation and therapy. Aliment. Pharmacol. Ther. 46 (2), 126–141.
Erker, J.C., Desai, S.M., Mushahwar, I.K., 1999. Rapid detection of Hepatitis E virus RNA by reverse transcription-polymerase chain reaction using universal oligonucleotide primers. J. Virol. Methods 81 (1–2), 109–113.
Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23 (2), 254–267.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874.
Lee, G.H., Tan, B.H., Teo, E.C., et al., 2016. Chronic infection with camelid hepatitis E virus in a liver transplant recipient who regularly consumes camel meat and milk. Gastroenterology 150 (2), 355–357e353.
Lu, L., Li, C., Hagedorn, C.H., 2006. Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. Rev. Med. Virol. 16 (1), 5–36.
Martin, D.P., Murrell, B., Golden, M., et al., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1 (1), vev003.
Meng, X.J., Purcell, R.H., Halbur, P.G., et al., 1997. A novel virus in swine is closely related to the human hepatitis E virus. Proc. Natl. Acad. Sci. U. S. A. 94 (18), 9860–9865.
Miura, M., Inoue, J., Tsuruoka, M., et al., 2017. Full-length genomic sequence analysis of new subtype 3k hepatitis E virus isolates with 99.97% nucleotide identity obtained from two consecutive acute hepatitis patients in a city in northeast Japan. J. Med. Virol. 89 (6), 1116–1120.
Monini, M., Di Bartolo, I., Ianiro, G., et al., 2015. Detection and molecular characterization of zoonotic viruses in swine fecal samples in Italian pig herds. Arch. Virol. 160 (10), 2547–2556.
Monne, I., Ceglie, L., Di Martino, G., et al., 2013. Hepatitis E virus genotype 4 in a pig farm, Italy. Epidemiol. Infect. 143 (3), 529–533.
Okamoto, H., 2011. Efficient cell culture systems for hepatitis E virus strains in feces and circulating blood. Rev. Med. Virol. 21 (1), 18–31.
Ricci, A., Allende, A., Bolton, D., et al., 2017. Public health risks associated with hepatitis E virus (HEV) as a food-borne pathogen. Efsa J. 15 (7).
Roth, A., Lin, J., Magnius, L., et al., 2016. Markers for ongoing or previous hepatitis E virus infection are as common in wild ungulates as in humans in Sweden. Viruses 8 (9).
Smith, D.B., Simmonds, P., Jameel, S., et al., 2014. Consensus proposals for classification of the family Hepeviridae. J. Gen. Virol. 95 (Pt 10), 2223–2232.
Smith, D.B., Simmonds, P., Izopet, J., et al., 2016. Proposed reference sequences for hepatitis E virus subtypes. J. Gen. Virol. 97 (3), 537–542.
Wang, Y., Ling, R., Erker, J.C., et al., 1999. A divergent genotype of hepatitis E virus in Chinese patients with acute hepatitis. J. Gen. Virol. 80 (Pt 1), 169–177.
Woo, P.C., Lau, S.K., Teng, J.L., et al., 2016. New hepatitis E virus genotype in bactrian camels, Xinjiang, China, 2013. Emerg. Infect. Dis. 22 (12), 2219–2221.

CHAPTER 4

Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy

4. Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy

Wild boars and pigs are considered the main reservoirs of the zoonotic genotypes of HEV (HEV-3 and HEV-4). Worldwide several studies described the presence of HEV in wild boars where several variants of HEV-3 were identified. HEV-4 was also rarely detected. The wide heterogeneity of HEV determines its classification in genotypes that are further divided in subtypes, named by letter (e.g. HEV-3e). In Europe including Italy, most of the detected wild boar strains have been classified into -3e, -3f and -3c subtypes and have been correlated to strains detected in humans and swine. The role of wild boar in HEV-3 transmission to humans is proved by several studies reporting the detection of HEV in wild boar liver and muscle and the occurrence of small outbreaks linked to consumption of undercooked wild boar meat. In some outbreaks, the foodborne transmission was directly proved by detection of the same HEV strain (100% nucleotide sequence identity) in food (raw wild boar meat) and patients. In Italy, no direct evidence of foodborne transmission has been described but some patients or subjects with IgG anti HEV have reported frequent consumption of wild boar meat or sausages or hunting activities. In Italy, demographic control of wild boar is mostly by hunting. Nevertheless, the wild boar population is constantly increasing. In this study, we reported the detection of HEV from 291 wild boars' livers in Southern Italy with 40 positive samples (13.7%) by Real Time RT-PCR. In order to characterize the detected HEV strains, a short region within the ORF2 was sequenced from 13 liver positive samples. The resulting phylogenetic analysis showed that several HEV-3 strains were circulating. Ten strains were classified into HEV-3c subtype that is the most common subtype infecting wild boar in Europe but is also frequently identified in human cases in France and in the Netherlands. Two strains were distant from the others and were provisionally classified in the HEV-3j subtype strains (HEV-3j like), due to low support of the phylogenetic analyses. The latter subtype has not been frequently detected and circulates rarely. The obtained results confirmed a heterogeneous population of HEV-3 strains in wild boar that is wider than in pigs. It is not clear if some HEV-3 strains circulate moderately or exclusively in wild boar. Probably, the lower density of wild boar than farmed pigs could partially justify the higher number of strains circulating in the former.

Received: 16 November 2016

DOI: 10.1111/tbed.12661

ORIGINAL ARTICLE

WILEY Transboundary and Emerging Diseases

Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy

G. Aprea[1,3] | M. G. Amoroso[1] | I. Di Bartolo[2] | N. D'Alessio[1] | D. Di Sabatino[3] | A. Boni[3] | B. Cioffi[1] | D. D'Angelantonio[3] | S. Scattolini[3] | L. De Sabato[2] | G. Cotturone[4] | F. Pomilio[3] | G. Migliorati[3] | G. Galiero[1] | G. Fusco[1]

[1]Istituto Zooprofilattico Sperimentale del Mezzogiorno, Portici, Italy

[2]Istituto Superiore di Sanità, Rome, Italy

[3]Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise "G. Caporale", Teramo, Italy

[4]Ente Parco Naturale Regionale Sirente Velino, Rocca di Mezzo, AQ, Italy

Correspondence
M. G. Amoroso, Istituto Zooprofilattico Sperimentale del Mezzogiorno, Portici, Italy.
Email: mamoroso@izsmportici.it

Summary

Hepatitis E virus (HEV) is a zoonotic pathogen with a worldwide distribution, and infects several mammalian species, including pigs and wild boars, which are recognized as its natural reservoirs. The virus causes a usually self-limiting liver disease with a mortality rate generally below 1%, although mortality rates of 15%–25% have been recorded in pregnant woman. Chronic infections can also occur. The prevalence of HEV has been extensively studied in wild boars and pigs in northern Italy, where intensive pig herds are predominantly located. In contrast, few data have been collected in south-central Italy, where small pig herds are surrounded by large regional parks populated with heterogeneous wild fauna. In this study, 291 liver samples from wild boars caught in south-central Italy were analysed with the molecular detection of viral RNA. Our results confirm the circulation of HEV in these animals, with a mean prevalence of 13.7% (40 of 291). A nucleotide sequence analysis showed that the HEV strains were highly conserved within the same geographic areas. The wild boar HEV strains belonged to the HEV-3c subtype, which is frequently described in wild boars, and to an uncommon undefined subtype (HEV-3j-like).The viral prevalence detected is concerning because it could represent a potential risk to hunters, meat workers and consumers of wild boar liver and derivative products. The hypothesized inter-species transmission of HEV to pigs and the possibility that the virus maintains its virulence in the environment and the meat chain also present potential risks to human health, and warrant further investigations in the near future.

KEYWORDS
Hepatitis E virus, wild boar, zoonosis

1 | INTRODUCTION

Hepatitis E is a liver disease caused by the Hepatitis E virus (HEV), a non-enveloped positive-sense single-stranded RNA virus classified in the family *Hepeviridae* and the genus *Orthohepevirus*. The *Orthohepevirus* A species include four major genotypes (HEV-1 to HEV-4) and infect humans, domestic pigs, wild boars and deers. Based on sequence analyses, these genotypes are further divided into subtypes. HEV-1 and HEV-2 cause infections in humans in developing countries, where they are transmitted primarily by the faecal–oral route through the consumption of contaminated water. Genotypes 3 and 4 are classified into 10 (a–j) and seven (a–g) subtypes (Lu, Li, & Hagedorn, 2006), respectively. HEV-3 is responsible for sporadic human infections in industrialized countries and can also infect

animals such as pigs and wild boars. There is a strict genetic correlation between the human and animal strains circulating in the same geographic area (Ruggeri et al., 2013). Recently, a case of HEV-3 in a pet rabbit was reported and the virus correlated strongly with a French genotype (Caruso, Modesto, Prato et al., 2015). The HEV-4 strain is endemic in pigs in Asia, where sporadic human cases associated with this genotype have been described. HEV-4 has also recently been detected in Italy in both pigs (Monne et al., 2015) and humans (Garbuglia et al., 2013). Novel genotypes have recently been reported in pigs, with HEV-5 and HEV-6 detected in wild boars in Japan, and HEV-7 detected in camels and a patient who regularly consumed camel milk (Lee et al., 2016). The foodborne transmission of HEV-3 and HEV-4 in industrialized countries, which is linked to the ingestion of uncooked deer and wild boar meat and raw pork liver, has been widely described in the literature (Tei, Kitajima, Takahashi, & Mishiro, 2003; Renou, Roque-Afonso, & Pavio, 2014; Ruggeri et al., 2013; Matsuda, Okada, Takahashi, & Mishiro, 2003; Li et al., 2005; Yazaki et al., 2003). Therefore, the disease is recognized as an emerging zoonosis and the dispersal of HEV raises public health concerns, especially in countries in which wild animals are hunted and their meat consumed by humans. There are numerous populations of wild boars in Italy, and around 300,000–500,000 individuals are estimated to exist throughout the country (Monaco, Franzetti, Pedrotti, & Toso, 2003). Boars are hunted from October to January, and their meat and entrails are used for direct human consumption or to produce sausages and salami. Several studies, mainly undertaken in north-central Italy, have described the detection of HEV in wild and domestic animals. Two recent papers (Caruso, Modesto, Bertolini et al., 2015; Caruso, Modesto, Prato et al., 2015; Di Profio et al., 2016) identified HEV in wild boars in northwestern Italy, characterizing the identified strains as subtypes 3e, 3f and 3c. However, no study has examined the prevalence of HEV in the wild boar populations of southern Italy. We conducted a molecular analysis of 291 livers collected from wild boars during the 2015 hunting season, in parks in three south-central Italian regions (Abruzzo, Campania and Calabria) to contribute new data on the prevalence of HEV in the south of Italy and to better understand the subtypes circulating in the wild reservoir. The results of our study clarify the prevalence and genetic variability of the HEV strains circulating in wild boars in Italy. Our data should also facilitate future research activities to clarify the risks posed to human consumers by raw boar liver food products and the potential role of wild boars in maintaining the virus among pig farms by their direct contact with these domestic animals.

2 | MATERIALS AND METHODS

2.1 | Sample collection

Two hundred ninety-one liver samples from wild boars (*Sus scrofa scrofa*) living in the parks and surrounding areas in three regions of south-central Italy were collected during the 2015–2016 hunting season (from October to January): from 144 animals in Abruzzo, 88 in Campania and 59 in Calabria. The animals were captured with selective mobile traps, and the necropsy procedures were performed in local slaughterhouses under the supervision of official veterinarians.

2.2 | Nucleic acid extraction

Nucleic acids were extracted with the MagMAX™ Express Instrument (Applied Biosystems, Monza, Italy). Liver samples (25 mg) were suspended in 2 ml of phosphate-buffered saline and homogenized with glass beads in TissueLyser (Qiagen, Milan, Italy). The samples were clarified by centrifugation, and 300 μl of the supernatant was loaded onto the MagMAX™ extraction sample plate. The RNA was extracted according to the manufacturer's instructions. The nucleic acids were eluted in 90 μl of elution buffer containing 40 μl of RNase inhibitor (Promega, Milan, Italy) and analysed immediately with real-time reverse transcription (RT)–PCR or stored at −80°C until use.

2.3 | Real-time RT–PCR detection of HEV

The primers and probes used in the real-time RT–PCR (Jothikumar, Cromeans, Robertson, Meng, & Hill, 2006) were as follows: forward primer HEV-F (5′-GGTGGTTTCTGGGGTGAC-3′), reverse primer HEV-R (5′-AGGGGTTGGTTGGATGAA-3′) and HEV probe (TaqMan probe) HEV-P (5′-FAM-TGATTCTCAGCCCTTCGC-BGQ1-3′). An internal amplification control (IAC) was also included in each reaction, using an MGB TaqMan probe: IACP (5′-VIC-CCATACACA TAGGTCAGG-MGB-NFQ-3′) (Martinez-Martinez, Diez-Valcarce, Hernandez, & Rodriguez-Lazaro, 2011). All the PCRs were performed on a 7500 Fast Real-Time PCR System (Applied Biosystems) with the following thermal profile: 50°C for 15 min, 95°C for 2 min, 45 cycles of 95°C for 10 s, 55°C for 20 s and 72°C for 15 s. The reaction mix (25 μl) contained 5 μl of extracted sample, 1× RNA UltraSense master mix (RNA UltraSense One-Step Quantitative RT–PCR System, Thermo Fisher Scientific, Rodano, Italy), 20× RNA UltraSense Enzyme Mix, 250 nM each primer (Tema Ricerca, Castesano, Italy), 100 nM probe (Thermo Fisher Scientific), 0.6 μl of IAC (Martinez-Martinez et al., 2011) and 50 nM IAC probe. The samples were considered positive when one replicate of the three tested showed a cycle threshold (Ct) value <38. Samples with Ct ≥ 38 were deemed negative or inhibited, depending on the result for IAC. The sample was considered negative if the Ct value for IAC was similar to the Ct value for IAC in the negative template controls (NTCs). If the sample IAC Ct value was higher than the NTC IAC Ct value, the sample was considered inhibited, and a 10-fold dilution of the sample was retested (Di Bartolo et al., 2012). The HEV-positive control strain was obtained from the Federal Research Institute for Animal Health (Germany).

2.4 | Phylogenetic analysis

RNA from the samples that was positive on real-time RT–PCR was analysed with the OneStep RT–PCR Kit (Qiagen), followed by a

second PCR amplification (nested PCR) with the Taq PCR Master Mix Kit (Qiagen), according to the manufacturer's instructions. The first set of primers used for the RT–PCR was 3156 (forward 5′-AAT(C)TATGCC(A)CAGTACCGGGTTG-3′) and 3157 (reverse 5′-CCCTTATCCTGCTGAGCATTCTC-3′); and the second set of primers for the nested PCR was 3158 (forward 5′-GTT(C)ATGC(T)TT(C)TGCATACATGGCT-3′) and 3159 (reverse 3′-AGCCGACGAAATC(T)AATTCTGTC-3′) (Huang et al., 2002). The RT–PCR and nested PCR were conducted according to the manufacturer's instructions, using an annealing temperature of 55°C. The 348-bp amplicons, amplified from within open reading frame 2 (ORF2), were sequenced and analysed as previously described (Amoroso et al., 2013; Di Bartolo et al., 2017). Nucleotide sequence similarity was analysed with the BLAST server (http://www.ncbi.nlm.nih.gov/genbank/index.html). A neighbour-joining (NJ) phylogenetic tree was constructed with the Tamura three-parameter model for nucleotides, using the MEGA 7 software (http://www.megasoftware.net/). Confidence values at the nodes were calculated with 1000 bootstrap replicates. The sequences were submitted to NCBI GenBank under accession numbers: WB/HEV/Ter03 KX549298, WB/HEV/Ter04 KX549299, WB/HEV/Ter05 KX549300, WB/HEV/Ter06 KX549301, WB/HEV/Ter07 KX549302, WB/HEV/Ter09 KX549303, WB/HEV/Ter10 KX549304, WB/HEV/Ter11 KX549305, WB/HEV/Ter12 KX549306, WB/HEV/NA18ITA15 KX549307, WB/HEV/NA19ITA15 KX549308, WB/HEV/NA20ITA15 KX549309 and WB/HEV/NA21ITA15 KX549310.

2.5 | Statistical analysis

For each region, the prevalence (P) and 95% confidence interval (CI) were calculated from the beta distribution with a Bayesian approach. A chi-square test was used to evaluate the probability that the difference in the prevalence of infection was attributable to chance. A p value <.05 was considered significant.

3 | RESULTS

In this study, we detected 40 HEV-positive wild boar livers among the 291 tested (P = 14%; 95% CI: 10.3–18.2). The prevalence differed according to the geographic origin of the animals tested (p = .001). In Abruzzo, HEV RNA was detected in 29 of 144 samples (20.14%; 95% CI: 14.4–27.4) (Figure 1), whereas in Campania, the prevalence was slightly lower (12.36%; 95% CI: 7.2–21.0), with 11 of the 88 samples positive on real-time RT–PCR (Figure 1). However, none of the samples from the regional parks of Calabria were positive for HEV (maximum expected prevalence 5%). The Ct values ranged from 24 to 37.8: in three samples, Ct < 28; in 10 samples, 28 ≤ Ct ≤ 34; and in 27 samples, Ct > 34. All the HEV-positive samples were also tested with nested RT–PCR before sequencing. The samples with Ct > 34 were negative according to the end-point RT–PCR and therefore were not included in the phylogenetic analysis. Overall, nine samples from Abruzzo and four samples from Campania were positive on the end-point RT–PCR and were sequenced. BLAST searches confirmed that all the HEV strains belonged to genotype 3. A phylogenetic tree (Figure 2) showed that eight of the nine HEV strains from the regional parks of Abruzzo were identical, whereas one strain shared 94.8% nucleotide identity with the other strains (accession number KX549306). The nine HEV strains from Abruzzo belonged to HEV-3c, clustering tightly with the prototype strain of the HEV-3c subtype (accession number FJ705359). Two of the four HEV strains detected in Campania also belonged to subtype HEV-3c, sharing 89.1%–90.6% nucleotide identity with the HEV-3c strains from Abruzzo. The remaining two strains from Campania (accession numbers KX549307 and KX549310) were not clearly assigned to any known subtype. They clustered with the prototype strain of subtype HEV-3j (accession number AY115488), but the cluster was poorly supported (bootstrap value, 27%) and the sequences shared only 88% nucleotide identity (Figure 2). An alignment of the deduced amino acid sequences showed that all the nucleotide changes occurred in the third base positions of codons, resulting in 100% amino acid identity within the same geographic area (data not shown).

4 | DISCUSSION

In this study, we investigated the presence of HEV in 291 liver samples from wild boars caught in parks in three regions of south-central Italy. The data on the occurrence of HEV in wild boars in Italy vary, ranging between 1.9% in northwestern Italy (Serracca et al., 2015) and 33.5% in central Italy (Montagnaro et al., 2015). The mean prevalence of HEV observed in our study was 14% (95% CI: 10.3–18.2), consistent with the previous finding of a greater presence of HEV in wild boars in south-central Italy than in those in the north (Di Profio et al., 2016). However, our results are difficult to compare with data already published because of the type of sampling strategy used in our study (for convenience), differences in the types of samples tested (faeces or liver), and because the HEV status of wild boars has never been investigated in two of the three regions examined here (Campania and Calabria). The significant differences in the prevalence of HEV related to the regions of the animals tested were mainly attributable to the Calabria data (chi-square test). Within the three regions investigated, the prevalence of HEV did not differ noticeably between Campania and Abruzzo. However, the failure to detect HEV in animals from regional parks in Calabria was unexpected. This result could be attributable to the poor representativeness of the animals tested (n = 59), resulting from the lack of a specific sampling design or to the actual low circulation of HEV in that area. In this context, the only study conducted on HEV in this region was performed in swine, with a described prevalence of 7.4% (Costanzo et al., 2015), lower than the mean prevalence reported for Italian swine.

Even though we tested opportunistically collected samples, the wild boars were present throughout the parks, and their distributions were contiguous with those in the surrounding areas, so we do not believe that our samples were unrepresentative. In a future study, a proper sampling strategy would make it possible to extend our findings from the parks to regional territories.

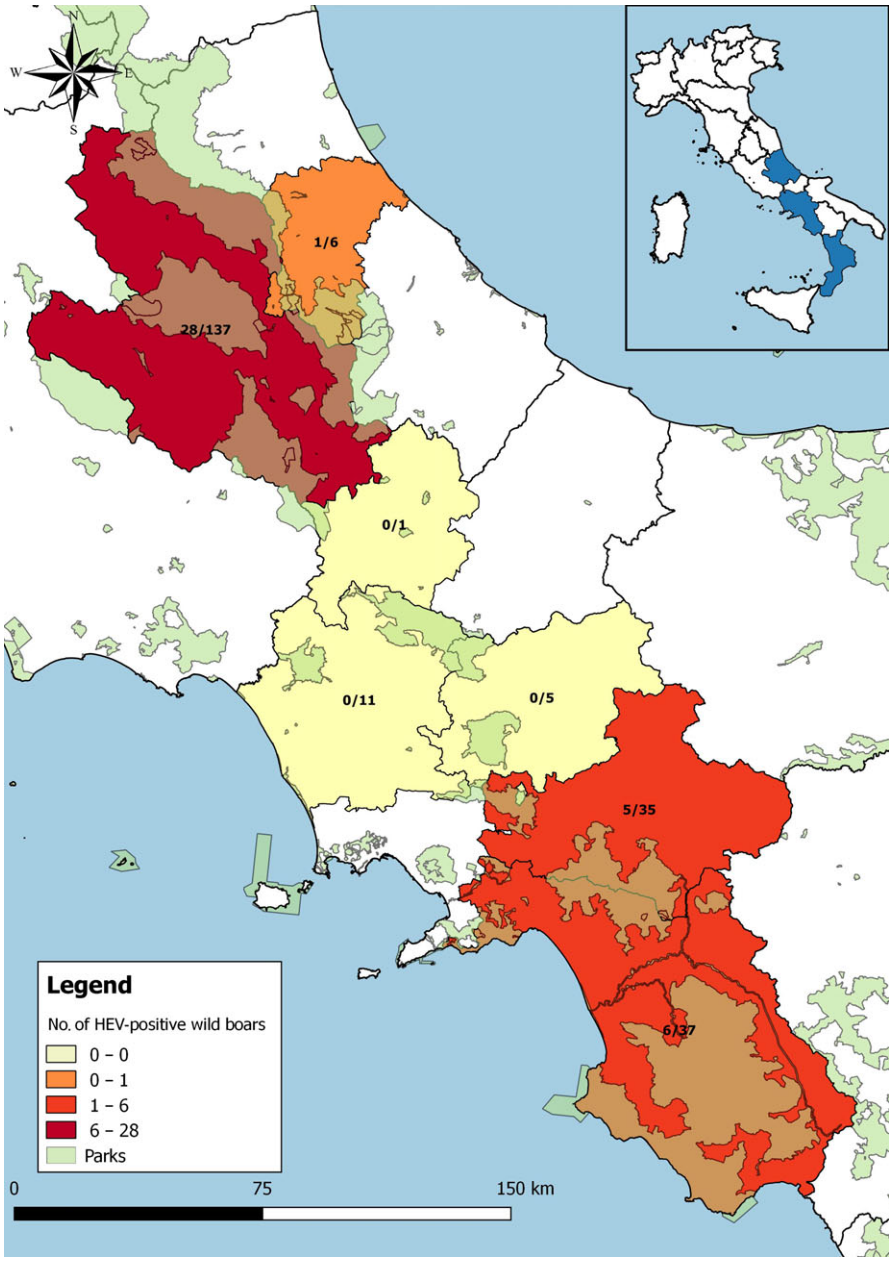


FIGURE 1 Geographical distribution of HEV-positive wild boars. The map shows the provinces of the south-central part of Italy in which positive samples were detected. Colours indicate different ranges of positivity as specified in the legend. For each province, the number of positive wild boars among the total analysed is reported. In the miniature in blue the regions investigated. A: Abruzzo; Ca: Campania; Cl: Calabria

Only 13 HEV-positive liver samples, with Ct ≤ 34 on real-time RT–PCR, were also positive on end-point RT–PCR. Our real-time RT–PCR, in which both the primers and probe targeted a conserved region in ORF2/3, may have been more sensitive than the end-point RT–PCR assay, which targeted a region within ORF2 (Gerber, Xiao, Cao, Meng, & Opriessnig, 2014).

Our phylogenetic analysis showed that the HEV strains from Abruzzo were identical, suggesting that only one strain was probably circulating among these wild boars at the time of sampling and that the introduction of new HEV strains or new animals from the surrounding areas is probably infrequent. In contrast, the samples from Campania grouped into two different clusters. Two HEV strains were classified as HEV-3c, whereas the other two strains were genetically distant from them and were not clearly classifiable. The two clusters contained HEV strains from two distinct geographic areas within the same region and the distance between the areas, or the presence of other geographic barriers, could explain the different HEV strains circulating in the two boar populations.

HEV-3c has frequently been described in wild boars and pigs in Europe (Caruso et al., 2016). The HEV-3c strain detected in Abruzzo shared 88%–96.4% nucleotide identity with Italian and German HEV-3c strains detected in wild boars (Di Profio et al., 2016; Schielke et al., 2009). Our phylogenetic analysis also showed that the HEV strains detected in the wild boar were genetically related to swine and human HEV strains circulating in Italy (Monini et al., 2015; Garbuglia et al., 2015). An HEV-3c strain was also detected in a patient with acute hepatitis, who was probably infected by the consumption of uncooked figatelli produced from pig liver (Garbuglia et al., 2015). The foodborne transmission of HEV from animal products to humans is an emerging concern and has been described in

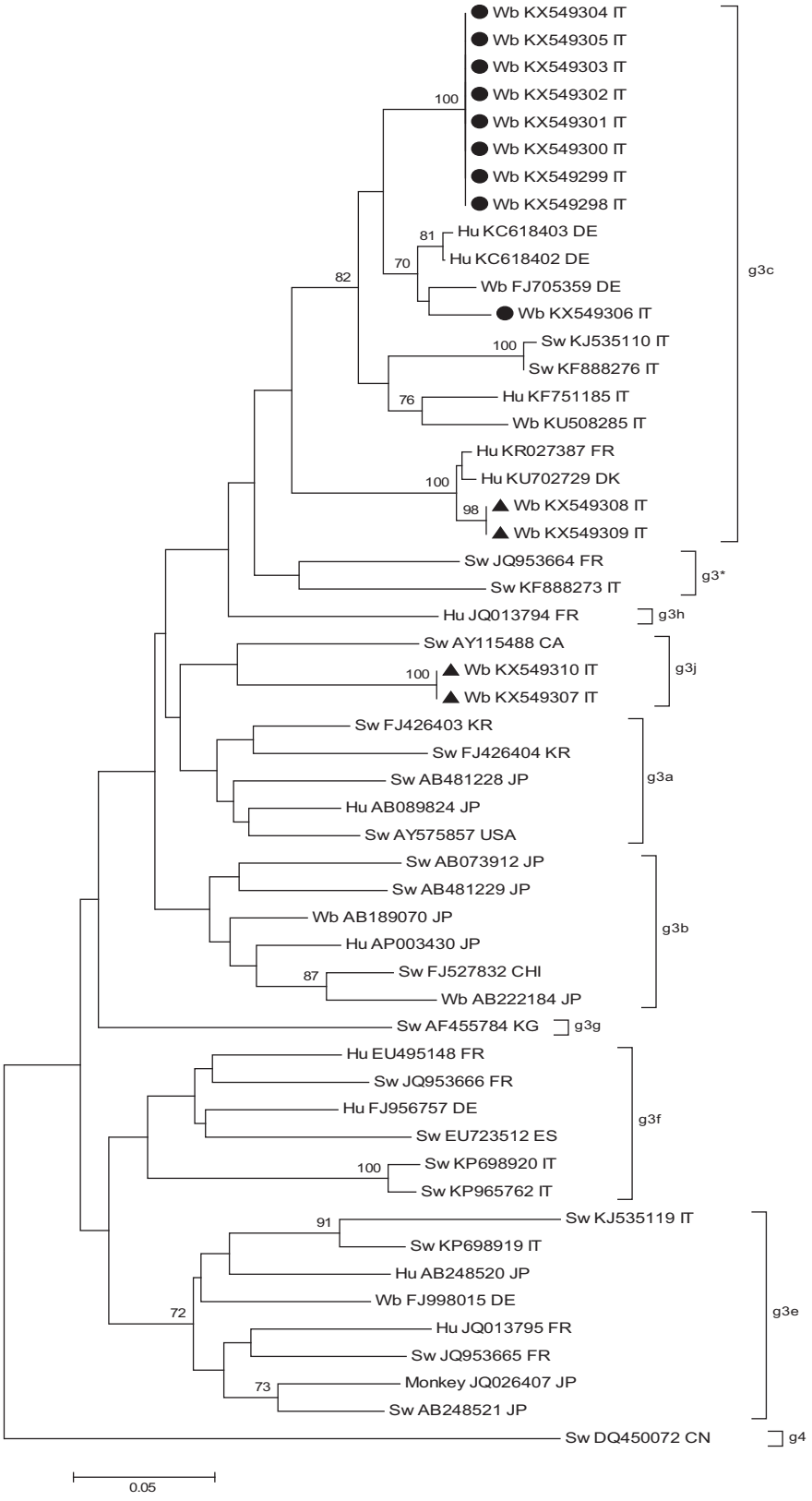


FIGURE 2 Phylogenetic tree based on 348 bp fragment of the ORF2 gene. The tree was generated by neighbour-joining using Tamura 3-parameter model and Swine g4 sequence as outgroup. Bootstraps values >70 are reported, obtained by supplying statistical support with bootstrapping of 1000 replicates. Sequences detected in this study are indicated by black dots (Abruzzo strains) and black triangles (Campania strains). Representative sequences of HEV genotype 3 strains from animals and humans have been included in the tree

several studies. Human infections have been reported after the consumption of contaminated meat products, particularly fresh liver sausages, which are commonly produced in Italy from both pig and boar liver (Colson et al., 2010, 2012). HEV genomes were also detected in sausages in a market (Di Bartolo, Angeloni, Ponterio, Ostanello, & Ruggeri, 2015). Strong evidence for the transmission of HEV-3 from wild boars to humans and human seroconversion was also documented in Italy in 2012 in people after wild boar butchering

(Giordani, Fabris, Brunetti, Goblirsch, & Romano, 2013). Possible interactions between domestic and wild animals could explain the maintenance of the virus in the environment (Wu et al., 2011). Most of the pig breeding located in south-central Italy are very small (with sometimes only two animals per farm) and biosecurity measures are poor, and there is evidence that wild animals interact with domestic animals, especially when attracted by agriculture products harvested nearby. In Italy, wild boars are widespread (Caruso, Modesto, Prato et al., 2015), which supports the hypothesis that wild animals are a reservoir for the virus, allowing it to be maintained in the environment. The HEV-3 strains detected in our study corroborate the zoonotic potential of the HEV strains circulating among wild animals in south-central Italy. Future surveys are required to investigate the presence of HEV in pigs reared at the borders of natural areas, close to regions in which wild boars are known to be HEV positive. Such surveys, together with comparative genomic sequencing of HEV, will test the hypothesis that the virus is transmitted from wild animals to farm animals.

ACKNOWLEDGEMENTS

This research was partially supported by the European Union's Seventh Framework Programme for Research, Technological Development and Demonstration under Grant Agreement no. 278433 PREDEMICS.

REFERENCES

Amoroso, M. G., Corrado, F., De Carlo, E., Lucibelli, M. G., Martucciello, A., Guarino, A., & Galiero, G. (2013). Bubaline herpesvirus 1 associated with abortion in a Mediterranean water buffalo. *Research in Veterinary Science*, *94*, 813–816.

Caruso, C., Modesto, P., Bertolini, S., Peletto, S., Acutis, P. L., Dondo, A., … Masoero, L. (2015). Serological and virological survey of hepatitis E virus in wild boar populations in northwestern Italy: Detection of HEV subtypes 3e and 3f. *Archives of Virology*, *160*, 153–160.

Caruso, C., Modesto, P., Prato, R., Scaglione, F. E., De Marco, L., Bollo, E., … Peletto, S. (2015). Hepatitis E virus: First description in a pet house rabbit. A new transmission route for human? *Transboundary and Emerging Diseases*, *62*, 229–232.

Caruso, C., Peletto, S., Rosamilia, A., Modesto, P., Chiavacci, L., Sona, B., … Masoero, L. (2016). Hepatitis E virus: A cross-sectional serological and virological study in pigs and humans at zoonotic risk within a high-density pig farming area. *Transboundary and Emerging Diseases*. doi: 10.1111/tbed.12533.

Colson, P., Borentain, P., Queyriaux, B., Kaba, M., Moal, V., Gallian, P., … Gerolami, R. (2010). Pig liver sausage as a source of hepatitis E virus transmission to humans. *Journal of Infectious Diseases*, *202*, 825–834.

Colson, P., Romanet, P., Moal, V., Borentain, P., Purgus, R., Benezech, A., … Gerolami, R. (2012). Autochthonous infections with hepatitis E virus genotype 4, France. *Emerging Infectious Disease*, *18*, 1361–1364.

Costanzo, N., Sarno, E., Peretti, V., Ciambrone, L., Casalinuovo, F., & Santoro, A. (2015). Serological and molecular investigation of swine hepatitis E virus in pigs raised in southern Italy. *Journal of Food Protection*, *78*, 2099–2102.

Di Bartolo, I., Angeloni, G., Ponterio, E., Ostanello, F., & Ruggeri, F. M. (2015). Detection of hepatitis E virus in pork liver sausages. *International Journal of Food Microbiology*, *193*, 29–33.

Di Bartolo, I., Diez-Valcarce, M., Vasickova, P., Kralik, P., Hernandez, M., Angeloni, G., … Ruggeri, F. M. (2012). Hepatitis E virus in pork production chain in Czech Republic, Italy, and Spain, 2010. *Emerging Infectious Diseases*, *18*, 1282–1289.

Di Bartolo, I., Ponterio, E., Angeloni, G., Morandi, F., Ostanello, F., Nicoloso, S., & Ruggeri, F. M. (2017). Presence of Hepatitis E virus in a RED deer (*Cervus elaphus*) population in central Italy. *Transboundary and Emerging Diseases*, *64*, 137–143.

Di Profio, F., Melegari, I., Sarchese, V., Robetto, S., Marruchella, G., Bona, M. C., … Di Martino, B. (2016). Detection and genetic characterization of hepatitis E virus (HEV) genotype 3 subtype c in wild boars in Italy. *Archives of Virology*, *161*, 2829–2834.

Garbuglia, A. R., Alessandrini, A. I., Pavio, N., Tesse, S., Grignolo, S., Viscoli, C., … Capobianchi, M. R. (2015). Male patient with acute hepatitis E in Genoa, Italy: Figatelli (pork liver sausage) as probable source of the infection. *Clinical Microbiology & Infection*, *21*, e4–e6.

Garbuglia, A. R., Scognamiglio, P., Petrosillo, N., Mastroianni, C. M., Sordillo, P., Gentile, D., … Capobianchi, M. R. (2013). Hepatitis E virus genotype 4 outbreak, Italy, 2011. *Emerging Infectious Diseases*, *19*, 110–114.

Gerber, P. F., Xiao, C. T., Cao, D., Meng, X. J., & Opriessnig, T. (2014). Comparison of real-time reverse transcriptase PCR assays for detection of swine hepatitis E virus in fecal samples. *Journal of Clinical Microbiology*, *52*, 1045–1051.

Giordani, M. T., Fabris, P., Brunetti, E., Goblirsch, S., & Romano, L. (2013). Hepatitis E and lymphocytic leukemia in Man, Italy. *Emerging Infectious Disease*, *19*, 2054–2056.

Huang, F. F., Haqshenas, G., Guenette, D. K., Halbur, P. G., Schommer, S. K., Pierson, F. W., … Meng, X. J. (2002). Detection by reverse transcription-PCR and genetic characterization of field isolates of swine hepatitis E virus from pigs in different geographic regions of the United States. *Journal of Clinical Microbiology*, *40*, 1326–1332.

Jothikumar, N., Cromeans, T. L., Robertson, B. H., Meng, X. J., & Hill, V. R. (2006). A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. *Journal of Virological Methods*, *131*, 65–71.

Lee, G. H., Tan, B. H., Teo, E. C., Lim, S. G., Dan, Y. Y., Wee, A., … Teo, C. G. (2016). Chronic infection with camelid hepatitis E virus in a liver transplant recipient who regularly consumes camel meat and milk. *Gastroenterology*, *150*(355–7), e3.

Li, T. C., Chijiwa, K., Sera, N., Ishibashi, T., Etoh, Y., Shinohara, Y., … Miyamura, T. (2005). Hepatitis E virus transmission from wild boar meat. *Emerging Infectious Diseases*, *11*, 1958–1960.

Lu, L., Li, C., & Hagedorn, C. H. (2006). Phylogenetic analysis of global hepatitis E virus sequences: Genetic diversity, subtypes and zoonosis. *Reviews in Medical Virology*, *16*, 5–36.

Martinez-Martinez, M., Diez-Valcarce, M., Hernandez, M., & Rodriguez-Lazaro, D. (2011). Design and application of nucleic acid standards for quantitative detection of enteric viruses by real-time PCR. *Food and Environmental Virology*, *3*, 92–98.

Matsuda, H., Okada, K., Takahashi, K., & Mishiro, S. (2003). Severe hepatitis E virus infection after ingestion of uncooked liver from a wild boar. *Journal of Infectious Diseases*, *188*, 944.

Monaco, A., Franzetti, B., Pedrotti, L., & Toso, S. (2003). *Linee Guida Per La Gestione Del Cinghiale*. Bologna, Italy: Ministero delle Politiche Agricole e Forestali. Istituto Nazionale Fauna Selvatica, Ozzano dell'Emilia.

Monini, M., Di Bartolo, I., Ianiro, G., Angeloni, G., Magistrali, C. F., Ostanello, F., & Ruggeri, F. M. (2015). Detection and molecular characterization of zoonotic viruses in swine fecal samples in Italian pig herds. *Archives of Virology*, *160*, 2547–2556.

Monne, I., Ceglie, L., DI Martino, G., Natale, A., Zamprogna, S., Morreale, A., … Bonfanti, L. (2015). Hepatitis E virus genotype 4 in a pig farm, Italy, 2013. *Epidemiology and Infection*, *143*, 529–533.

Montagnaro, S., De Martinis, C., Sasso, S., Ciarcia, R., Damiano, S., Auletta, L., ... Pagnini, U. (2015). Viral and antibody prevalence of hepatitis E in European wild boars (*Sus scrofa*) and hunters at zoonotic risk in the Latium region. *Journal of Comparative Pathology*, *153*, 1–8.

Renou, C., Roque-Afonso, A. M., & Pavio, N. (2014). Foodborne transmission of hepatitis E virus from raw pork liver sausage, France. *Emerging Infectious Disease*, *20*, 1945–1947.

Ruggeri, F. M., Di Bartolo, I., Ponterio, E., Angeloni, G., Trevisani, M., & Ostanello, F. (2013). Zoonotic transmission of hepatitis E virus in industrialized countries. *New Microbiologica*, *36*, 331–344.

Schielke, A., Sachs, K., Lierz, M., Appel, B., Jansen, A., & Johne, R. (2009). Detection of hepatitis E virus in wild boars of rural and urban regions in Germany and whole genome characterization of an endemic strain. *Virology Journal*, *6*, 58.

Serracca, L., Battistini, R., Rossini, I., Mignone, W., Peletto, S., Boin, C., ... Ercolini, C. (2015). Molecular investigation on the presence of hepatitis E virus (HEV) in wild game in north-western Italy. *Food and Environmental Virology*, *7*, 206–212.

Tei, S., Kitajima, N., Takahashi, K., & Mishiro, S. (2003). Zoonotic transmission of hepatitis E virus from deer to human beings. *Lancet*, *362*, 371–373.

Wu, N., Abril, C., Hinic, V., Brodard, I., Thur, B., Fattebert, J., ... Ryser-Degiorgis, M. P. (2011). Free-ranging wild boar: A disease threat to domestic pigs in Switzerland? *Journal of Wildlife Diseases*, *47*, 868–879.

Yazaki, Y., Mizuo, H., Takahashi, M., Nishizawa, T., Sasaki, N., Gotanda, Y., & Okamoto, H. (2003). Sporadic acute or fulminant hepatitis E in Hokkaido, Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. *Journal of General Virology*, *84*, 2351–2357.

How to cite this article: Aprea G, Amoroso MG, Di Bartolo I, et al. Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy. *Transbound Emerg Dis*. 2018;65:e25–e31. https://doi.org/10.1111/tbed.12661

CHAPTER 5

Phylogenetic analysis of two genotype 3 Hepatitis E viruses from wild boar, Italy

5. Phylogenetic analysis of two genotype 3 Hepatitis E viruses from wild boar, Italy

In the previous Chapter, we described the detection of HEV strains from a wild boar population hunted in South Italy. Among HEV-3 strains detected by Real-time RT-PCR, one of the strain was not amplified by end-point RT-PCR amplifying short conserved genome regions. We supposed that the amplification did not succeed for low identity between strain sequences and the primers. This strain, together with another HEV-3 provisionally classified using short genome region as HEV-3-j-like, were subjected to NGS (method described in the first Chapter). Full genomes were obtained for both strains and subjected to phylogenetic analyses. The strain, that was previously classified as HEV-3-j-like in a not statistically supported phylogenetic tree, after the analysis of the complete genome, clustered within the HEV-3i subtype reference sequences. This is the first identification of the HEV-3i in Italy. This subtype has been rarely detected in humans and has been described only recently in wild animals (wild boar and roe deer) in Lituania. The other sequenced strain showed evolutionary correlation with a human strain identified in Japan but not classified in any subtypes because distant from the reference strains, probably representing a novel subtype. Furthermore, sequence analysis revealed a limited number of matching with this strain that could reasonably circulate moderately. The reason why some subtypes are less frequent is unknown maybe it could be linked to recent adaption in wild animals.

Virus Genes
https://doi.org/10.1007/s11262-018-1597-5

Phylogenetic analysis of two genotype 3 Hepatitis E viruses from wild boar, Italy

Luca De Sabato[1,2] · Gabriele Vaccari[1] · Philippe Lemey[3] · Maria Grazia Amoroso[4] · Giovanna Fusco[4] · Giovanni Ianiro[1] · Ilaria Di Bartolo[1]

Received: 8 May 2018 / Accepted: 31 August 2018

Abstract

The complete and near-complete genome sequences (7206 nt and 7229 nt) of two wild boar HEV strains detected in Southern Italy were obtained by the next generation sequencing. Phylogenetic analysis and *p* distance comparisons of one of the strains with HEV-3 reference subtype strains confirmed the detection of a subtype 3i (*p* distance = 0.110) strain in wild boar, never detected in Italy either in wild boar or pigs. The sequence of the second strain was not classifiable in any of the subtypes defined to date, showing a *p* distance > 0.138 and a low nucleotide identity with all HEV reference strains. The virus may represent a novel subtype, with a low relationship to other strains of genotype 3 detected in wild boar, pigs, or humans in Europe. This result suggests the circulation in Italy of an emerging or uncommon HEV strain. Sequencing followed by phylogenetic analyses of the complete HEV coding regions are important tools for understanding the evolutionary and epidemiological dynamics underlying the wide genetic diversity of HEV strains.

Keywords Hepatitis E virus · Wild boar · Full genome · Next generation sequencing · Subtype · HEV-3

Hepatitis E is an acute self-limiting liver disease, caused by the Hepatitis E virus (HEV) [1]. The viral genome (ssRNA) is structured in short 5′ and 3′ untranslated regions (UTR) and three Open Reading Frames (ORF1, ORF2, and ORF3) encoding: the non-structural proteins, the capsid protein, and a small multifunctional phosphoprotein [2], respectively. HEV includes strains infecting humans, pigs, wild boar, deer, mongooses, rabbits, and camels, which are classified into seven genotypes (HEV1–7) with the additionally proposed genotype 8, recently detected in Bactrian camels [3]. The HEV genotypes are divided into subtypes based on the nucleotide sequence difference of HEV complete genomes [4–7]. A novel subtype can be established if at least three genomes, from epidemiologically unrelated strains forming a cluster divergent from existing subtypes, are available [6].

HEV-3 and HEV-4 are zoonotic and in Europe the major route of transmission is foodborne. Small outbreaks and sporadic cases have been associated with the consumption of raw or undercooked meat from wild boar and deer or pork liver sausages. Pigs and wild boar are considered to be the main HEV reservoirs [8]. In Europe, the reported prevalence of HEV-3 in wild boar ranges from 3.7% [9] to 33.5% [9] and the strains detected showed a high genetic diversity [6]. In Italy, several studies have reported the detection of HEV in wild boar, and, based on short genome regions, the strains have been assigned to HEV-3c, -3e, and -3f subtypes [9–16]. In a recent study conducted on wild boar in two regional parks in Southern Italy (Campania), two wild boar HEV strains (namely WB/HEV/NA17ITA15 and WB/HEV/NA21ITA15) were detected by Real-Time RT-PCR. For the WB/HEV/NA17ITA15 strain no sequence data were available. Sequence analyses of the short genome region in the

Edited by William Dundon.

✉ Gabriele Vaccari
gabriele.vaccari@iss.it

[1] Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore di Sanità, Viale Regina Elena, 299, 00161 Rome, Italy

[2] Department of Sciences, University Roma Tre, Viale Guglielmo Marconi 446, 00146 Rome, Italy

[3] Department of Microbiology and Immunology, Rega Institute, KU Leuven – University of Leuven, Herestraat 49, 3000 Louvain, Belgium

[4] Department of Animal Health, Istituto Zooprofilattico del Mezzogiorno, Via Salute, 2, 80055 Portici, NA, Italy

Published online: 10 September 2018

ORF2 of WB/HEV/NA21ITA15 did not allow for definitive classification with any HEV subtypes defined to date [16]. The present study aimed to characterize the two strains (WB/HEV/NA17ITA15 and WB/HEV/NA21ITA15) and to clarify their subtype classification based on the full coding region genome sequencing obtained by next generation sequencing (NGS).

Total RNA was extracted from 250 mg of wild boar liver samples with the RNeasy Mini Kit (Qiagen, Milan, Italy) and quantified by real-time quantitative reverse-transcription PCR (RT-qPCR), as previously described [17]. Total RNA was subjected to the sequence independent single primer amplification (SISPA) [18], and the library was obtained by the NEBNext Fast DNA Library Prep Set for Ion Torrent (New England BioLabs, Hitchin, United Kingdom) following the manufacturers' protocol and sequenced on an Ion Personal Genome Machine (PGM) in Ion 318 Chip v2 (Thermo Fisher Scientific, Rodano, Italy). The reads obtained by the NGS run were analyzed by Galaxy Aries (https://aries.iss.it) as previously described [19].

Two contigs of 500 and 6500 bases for WB/HEV/NA17ITA15 and a contig of 7200 nt long for WB/HEV/NA21ITA15 were obtained by de novo assembly.

The genome of the WB/HEV/NA17ITA15 was completed by dideoxy chain-termination method, using a semi-nested reverse-transcription PCR (RT-PCR) to amplify a 500 bp fragment in the ORF1 (methyltransferase), using one primer designed at the 5′-UTR (5′-CCACRTATGTGGTCGATG CC-3′) and two primers designed at the contig sequences: HEV558Rw (5′-AGGTCGTGCAAAGAATAAAG-3′) and HEV544Rw (5′-ATAAAGGGCTATGCCAGTCT-3′). Full genomes were submitted to GenBank under the accession numbers: MF959764–MF959765.

The genome organization was established predicting the ORFs using the online tool ORF finder (NCBI, http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Recombination events were tested applying the Phi test [20] implemented in SplitsTree version 4 software.

Maximum likelihood (ML) phylogenetic analyses using the general time reversible model with gamma distribution plus invariant sites (GTR + G + I) suggested by model-test and *p* distance calculation were performed using MEGA7 [21]. Support values for the clusters were obtained by bootstrap analyses using 1000 pseudo-replicates.

The HEV reads were obtained by sequencing 3.8×10^6 genome equivalents (GE) for WB/HEV/NA17ITA15 and 1.7×10^8 GE for WB/HEV/NA21ITA15. The two genomes WB/HEV/NA17ITA15 and WB/HEV/NA21ITA15 were 7206 nt and 7229 nt long and organized in the three typical HEV ORFs. We did not succeed in sequencing the 5′ end of the genome of WB/HEV/NA17ITA15. The Italian strains were compared with 80 HEV genotype 3 complete genomes (lacking the 5' UTR), including the reference genomes suggested by Smith et al. [6] and full genome sequences that displayed the highest nucleotide similarity with the strains sequenced in this study. WB/HEV/NA17ITA15 showed the highest nucleotide identity (89.0% ni.), and amino acid identity (ORF1: 88.2% ni 97.5% aa; ORF2: 89.5% ni 98.4% aa; ORF3: 95.9% ni 94.3% aa) with the 3i prototype (FJ998008, BB02) [6]. The WB/HEV/NA17ITA15 strain clustered with the 3i subtype prototype (*p* distance = 0.110) and showed a *p* distance value > 0.128 for all the other HEV subtype reference strains and was definitively assigned to the 3i subtype. The WB/HEV/NA21ITA15 showed 86.2% ni with two unclassified HEV-3 strains: IC2011 (KU513561) and MWP_2010 (KP294371). However, considering the single ORFs identities, strain WB/HEV/NA21ITA15 displayed the highest amino acid similarity in the ORF1 with the 3 h (97.2%), 3i (96.8%), and 3c (96.9%) subtypes prototypes, while the highest amino acid similarity in ORF2 (97.3%) and ORF3 (97.7%) was with IC2011 (unassigned) and 3 h prototype strain. Lower amino acid identities were observed with the 3i subtype prototype strain for both ORF2 (97.0%) and ORF3 (91.2%).

After excluding any evidence of recombination ($p = 0.98$), phylogenetic reconstruction was performed for the dataset described above.

In the phylogenetic tree (Fig. 1), the WB/HEV/NA21ITA15 formed a cluster with the Japanese unclassifiable human strain IC2011 (KU513561), included in the clade of the 3chi subtypes prototypes together with some additional unassigned strains. Neither the ML tree bootstrap values (34–42) for the clustering of WB/HEV/NA21ITA15 nor the *p* distance > 0.138 to all prototypes of HEV subtypes supported a particular subtype assignment (Fig. 1).

We further investigated the amino acid differences of WB/HEV/NA21ITA15 by comparing its strain amino acid sequence with the sequence dataset (82 full genomes). Results revealed 12 unique amino acid changes not present in the other HEV-3 strains (Table 1). The three amino acid changes in the capsid protein were mapped to the S domain at the beginning of the N-terminal arginine-rich region. The additional two amino acid changes are in the M and P domains and were predicted to be in the α-helix secondary structure (data not shown, by online tool Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index). In the absence of protein structures in the database, it was not possible to map the other amino acid changes.

Interestingly, two amino acid substitutions in the helicase (D996E) and protease regions (H581Q) (ORF1) were shared with HEV-4 strains, as shown by comparisons of the WB/HEV/NA21ITA15 with ten full genomes of HEV-4 (AB197673, DQ279091, AB074915, AJ272108, AY723745, AB220974, AB108537, GU119961, DQ450072, AB369688).

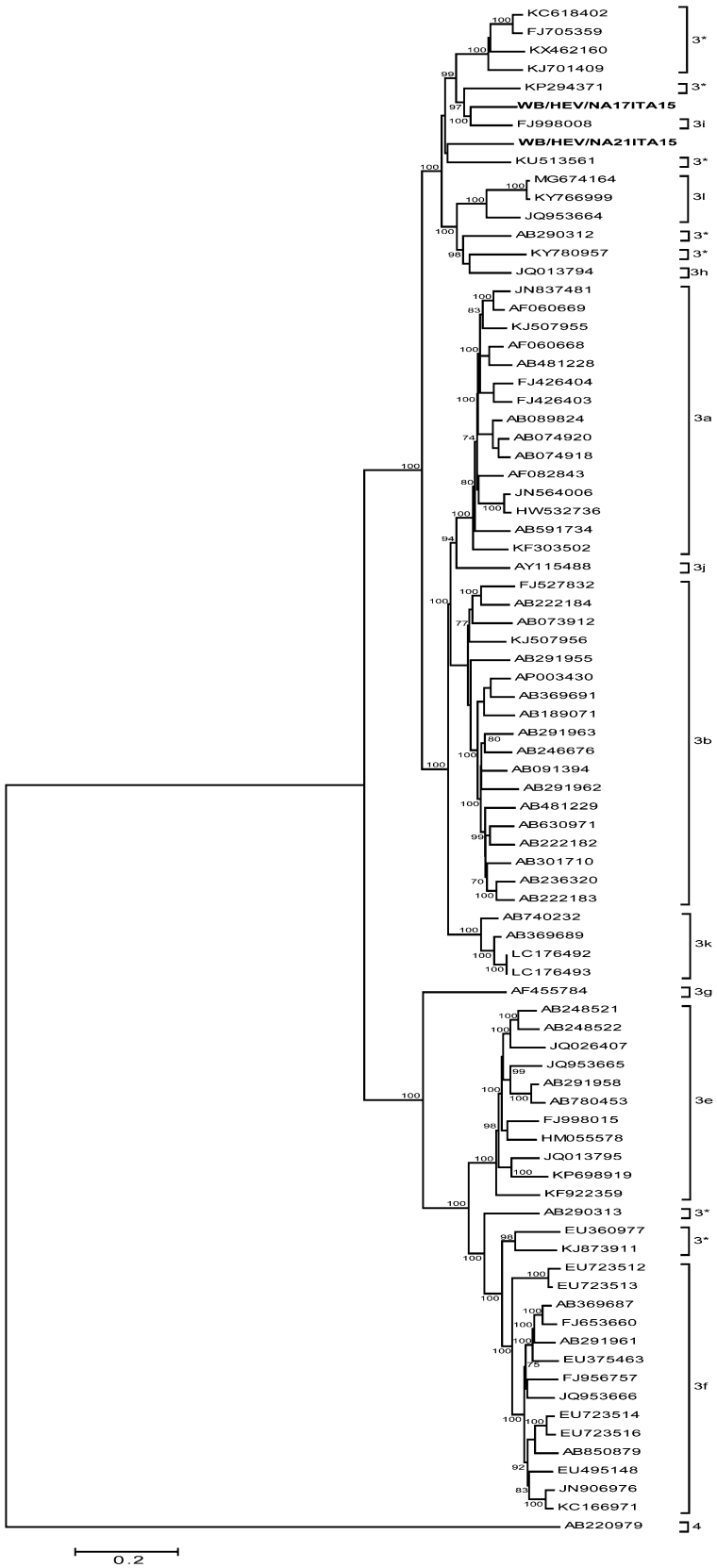


Fig. 1 The maximum likelihood phylogenetic tree was built with GTR + G + I substitution model, by 1000 resampling. The tree included 82 HEV-3 genotype complete genome sequences; Italian strains WB/HEV/NA17ITA15 and WB/HEV/NA21ITA15 are indicated in bold. HEV-4 strain as outgroup and bootstrap replicates > 70% were reported

Table 1 Polymorphic positions in HEV-3 reference sequences (PP%) and WB/HEV/NA21ITA15 sequence

Protein	ORF1 1703 aa							Capsid protein (ORF2) 645 aa						pORF3 113 aa	
	Met. ase[a]	Y domain	Polypro-line (vari-able)	Protease	X domain	Helicase	RdRp[b]	S domain (N-ter-minal)	S domain (C-termi-nal)	M domain	P domain	D1 domain	D2 domain	P1 domain	P2 domain
Summarized function	mRNA ribo-some bind-ing, innate immu-nity [28]	Not known	Adaptation [28] and immune response modula-tion [29]	Protease [30], deu-biquit-ination and cellular immu-nity [29]	Not known	Helicase [28]	Replica-tion [31]	Capsid [32]	Not known	Connect-ing S and P domains [32]	Neutral-izing anti-bodies binding and cellular receptor [28]	Cell sur-vival signal [33]	Innate immune response [34]	ORF2 interac-tion (ser71) [31]	Virus release [31]
Length (AA)	184	218	Variable	160	174	233	487	113	203	133	208	16	25	11	19
Amino acid exchanges	L49F D173E	No changes	No changes	H581Q	No changes	D996E A^M^S^G^L 1032V	No changes	R28G G31R R32G	No changes	L375I	T^S497A	P13T	V^A43D	No changes	No changes

[a]Methyltransferase

[b]RNA-dependent RNA polymerase

In the present study, we fully characterized two HEV strains identified in wild boar, one of which (WB/HEV/NA17ITA15) was classified as 3i, a subtype already detected in boars in Europe, but never detected before in Italy either in animals (pigs and wild boars) or in humans [22–26]. Interesting the second strain WB/HEV/NA21ITA15 could not be unambiguously classified with any of the subtypes defined to date.

Since the WB/HEV/NA21ITA15 displayed a low nucleotide sequence identity (< 89%) with the other HEV-3 strains when analyzing both full and partial genomes (ORF2, Met.asi and RdRp fragments commonly used for diagnoses, data not shown) available online, it can be considered uncommon. The amino acid changes in the S domain (R → G and vice versa) displayed by WB/HEV/NA21ITA15 (Table 1) are not surprising, since this region is generally rich in arginine and glycine. The two amino acid substitutions in the helicase (D996E) and protease regions (H581Q) were common with HEV-4 strains, but absent in the 82 HEV-3 genotype full genomes included in the comparison.

The amino acid mapping analyses cannot lead to any conclusions about differences in infectivity or host specificity, which is still difficult to assess. In addition, we did not succeed in cultivating WB/HEV/NA21ITA15 strain on cell culture (data not shown), not allowing for any further correlation between the amino acid substitutions detected and changes in the infectivity of this HEV viral variant. Until specific *p* distance cut-off values for HEV-3 subtype assignment are established, it will not be possible to assign full genomes to a subtype that do not show robust clustering with known reference sequences. These genomes could represent novel subtypes or variants of the already established subtypes.

The circulation of rare HEV strains deserves additional investigation in order to understand the role of wild boars in HEV circulation, and the risk of infection by newly emerging strains through interspecies or zoonotic transmissions.

NGS represents a powerful method to obtain full genomes, but it is still not widely applied to HEV.

It is notable that the WB/HEV/NA17ITA15 was positive by Real-time RT-PCR, but not by diagnostic RT-PCRs, and hence no sequence data were available [16]. The use of the SISPA coupled with NGS allowed us to overcome this limitation. However, the use of NGS for HEV sequencing appears to be restricted to particular types of samples. In this study, 3.8×10^6 to 1.7×10^8 GE were sufficient to sequence the full genome, whereas in our previous study conducted from feces a comparable amount of virus GE did not lead to the same result [27]. This may be due to a different percentage of viral nucleic acids present in the fecal samples.

In conclusion, in the absence of a clear cut-off value to classify HEV variants in subtypes, full genome sequences are needed to conduct comprehensive phylogenetic analyses and fully characterize circulating strains and subtypes.

Author Contributions LDS, GV, PL, and IDB conceived and designed the experiments; LDS and GI performed the experiments; GF, BC, and MGA contributed to diagnosis and collected samples. LDS, GV, PL, and IDB contributed to the writing and revision of the manuscript. All authors read and approved the final manuscript.

Funding This work was partially supported by the European Union's Seventh Framework Programme for research, technological development and demonstration under Grant Agreement No. 278433 PREDEMICS.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Research involving human and/or animal participants No animals were used.

References

1. Donnelly MC, Scobie L, Crossan CL, Dalton H, Hayes PC, Simpson KJ (2017) Review article: Hepatitis E—a concise review of virology, epidemiology, clinical presentation and therapy. Aliment Pharmacol Ther 46:126–141. https://doi.org/10.1111/apt.14109
2. Tam AW, Smith MM, Guerra ME, Huang CC, Bradley DW, Fry KE, Reyes GR (1991) Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. Virology 185:120–131
3. Woo PC, Lau SK, Teng JL, Cao KY, Wernery U, Schountz T, Chiu TH, Tsang AK, Wong PC, Wong EY, Yuen KY (2016) New Hepatitis E virus genotype in bactrian camels, Xinjiang, China, 2013. Emerg Infect Dis 22:2219–2221. https://doi.org/10.3201/eid2212.160979
4. Smith DB, Simmonds P, Jameel S, Emerson SU, Harrison TJ, Meng XJ, Okamoto H, Van der Poel WH, Purdy MA (2014) Consensus proposals for classification of the family Hepeviridae. J Gen Virol 95:2223–2232. https://doi.org/10.1099/vir.0.068429-0
5. Lee GH, Tan BH, Teo EC, Lim SG, Dan YY, Wee A, Aw PP, Zhu Y, Hibberd ML, Tan CK, Purdy MA, Teo CG (2016) Chronic infection with camelid Hepatitis E virus in a liver transplant recipient who regularly consumes camel meat and milk. Gastroenterology 150:355–357 e353. https://doi.org/10.1053/j.gastro.2015.10.048
6. Smith DB, Simmonds P, Izopet J, Oliveira-Filho EF, Ulrich RG, Johne R, Koenig M, Jameel S, Harrison TJ, Meng XJ, Okamoto H, Van der Poel WH, Purdy MA (2016) Proposed reference sequences for Hepatitis E virus subtypes. J Gen Virol 97:537–542. https://doi.org/10.1099/jgv.0.000393
7. Miura M, Inoue J, Tsuruoka M, Nishizawa T, Nagashima S, Takahashi M, Shimosegawa T, Okamoto H (2017) Full-length genomic sequence analysis of new subtype 3 k Hepatitis E virus isolates with 99.97% nucleotide identity obtained from two consecutive acute hepatitis patients in a city in northeast Japan. J Med Virol 89:1116–1120. https://doi.org/10.1002/jmv.24743
8. Adlhoch C, Avellon A, Baylis SA, Ciccaglione AR, Couturier E, de Sousa R, Epstein J, Ethelberg S, Faber M, Feher A, Ijaz S, Lange H, Mandakova Z, Mellou K, Mozalevskis A, Rimhanen-Finne R, Rizzi V, Said B, Sundqvist L, Thornton L, Tosti ME, van

Pelt W, Aspinall E, Domanovic D, Severi E, Takkinen J, Dalton HR (2016) Hepatitis E virus: assessment of the epidemiological situation in humans in Europe, 2014/15. J Clin Virol 82:9–16. https://doi.org/10.1016/j.jcv.2016.06.010

9. Montagnaro S, De Martinis C, Sasso S, Ciarcia R, Damiano S, Auletta L, Iovane V, Zottola T, Pagnini U (2015) Viral and antibody prevalence of Hepatitis E in European wild boars (*Sus scrofa*) and hunters at zoonotic risk in the Latium region. J Comp Pathol 153:1–8. https://doi.org/10.1016/j.jcpa.2015.04.006
10. Martelli F, Caprioli A, Zengarini M, Marata A, Fiegna C, Di Bartolo I, Ruggeri FM, Delogu M, Ostanello F (2008) Detection of Hepatitis E virus (HEV) in a demographic managed wild boar (*Sus scrofa* scrofa) population in Italy. Vet Microbiol 126:74–81. https://doi.org/10.1016/j.vetmic.2007.07.004
11. Caruso C, Modesto P, Bertolini S, Peletto S, Acutis PL, Dondo A, Robetto S, Mignone W, Orusa R, Ru G, Masoero L (2015) Serological and virological survey of Hepatitis E virus in wild boar populations in northwestern Italy: detection of HEV subtypes 3e and 3 f. Arch Virol 160:153–160. https://doi.org/10.1007/s00705-014-2246-5
12. Martinelli N, Pavoni E, Filogari D, Ferrari N, Chiari M, Canelli E, Lombardi G (2015) Hepatitis E virus in wild boar in the central northern part of Italy. Transbound Emerg Dis 62:217–222. https://doi.org/10.1111/tbed.12118
13. Mazzei M, Nardini R, Verin R, Forzan M, Poli A, Tolari F (2015) Serologic and molecular survey for Hepatitis E virus in wild boar (Sus scrofa) in Central Italy. New Microbes New Infect 7:41–47. https://doi.org/10.1016/j.nmni.2015.05.008
14. Serracca L, Battistini R, Rossini I, Mignone W, Peletto S, Boin C, Pistone G, Ercolini R, Ercolini C (2015) Molecular investigation on the presence of Hepatitis E virus (HEV) in wild game in North-Western Italy. Food Environ Virol 7:206–212. https://doi.org/10.1007/s12560-015-9201-9
15. Di Profio F, Melegari I, Sarchese V, Robetto S, Marruchella G, Bona MC, Orusa R, Martella V, Marsilio F, Di Martino B (2016) Detection and genetic characterization of Hepatitis E virus (HEV) genotype 3 subtype c in wild boars in Italy. Arch Virol 161:2829–2834. https://doi.org/10.1007/s00705-016-2964-y
16. Aprea G, Amoroso MG, Di Bartolo I, D'Alessio N, Di Sabatino D, Boni A, Cioffi B, D'Angelantonio D, Scattolini S, De Sabato L, Cotturone G, Pomilio F, Migliorati G, Galiero G, Fusco G (2017) Molecular detection and phylogenetic analysis of Hepatitis E virus strains circulating in wild boars in south-central Italy. Transbound Emerg Dis 65:e25–e31. https://doi.org/10.1111/tbed.12661
17. Di Bartolo I, Angeloni G, Monini M, Maione E, Marrone R, Ostanello F, Ruggeri FM (2016) Complete genome sequencing of a genotype 3 Hepatitis E virus strain identified in a swine farm in Italy. Virus Res 211:89–95. https://doi.org/10.1016/j.virusres.2015.09.014
18. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ (2008) Viral genome sequencing by random priming methods. BMC Genom 9:5. https://doi.org/10.1186/1471-2164-9-5
19. Moreno A, Lelli D, de Sabato L, Zaccaria G, Boni A, Sozzi E, Prosperi A, Lavazza A, Cella E, Castrucci MR, Ciccozzi M, Vaccari G (2017) Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy. Virol J 14:239. https://doi.org/10.1186/s12985-017-0907-1
20. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267. https://doi.org/10.1093/molbev/msj030
21. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054
22. Adlhoch C, Wolf A, Meisel H, Kaiser M, Ellerbrok H, Pauli G (2009) High HEV presence in four different wild boar populations in East and West Germany. Vet Microbiol 139:270–278. https://doi.org/10.1016/j.vetmic.2009.06.032
23. Schielke A, Sachs K, Lierz M, Appel B, Jansen A, Johne R (2009) Detection of Hepatitis E virus in wild boars of rural and urban regions in Germany and whole genome characterization of an endemic strain. Virol J 6:58. https://doi.org/10.1186/1743-422x-6-58
24. Oliveira-Filho EF, Bank-Wolf BR, Thiel HJ, Konig M (2014) Phylogenetic analysis of Hepatitis E virus in domestic swine and wild boar in Germany. Vet Microbiol 174:233–238. https://doi.org/10.1016/j.vetmic.2014.09.011
25. Vina-Rodriguez A, Schlosser J, Becher D, Kaden V, Groschup MH, Eiden M (2015) Hepatitis E virus genotype 3 diversity: phylogenetic analysis and presence of subtype 3b in wild boar in Europe. Viruses Basel 7:2704–2726. https://doi.org/10.3390/v7052704
26. Anheyer-Behmenburg HE, Szabo K, Schotte U, Binder A, Klein G, Johne R (2017) Hepatitis E virus in wild boars and spillover infection in red and roe deer, Germany, 2013–2015. Emerg Infect Dis 23:130–133. https://doi.org/10.3201/eid2301.161169
27. De Sabato L, Lemey P, Vrancken B, Bonfanti L, Ceglie L, Vaccari G, Di Bartolo I (2018) Proposal for a new subtype of the zoonotic genotype 3 Hepatitis E virus: HEV-3 l. Virus Res 248:1–4. https://doi.org/10.1016/j.virusres.2018.02.010
28. Holla RP, Ahmad I, Ahmad Z, Jameel S (2013) Molecular virology of hepatitis E virus. Semin Liver Dis 33(1):3–14. https://doi.org/10.1055/s-0033-1338110
29. Karpe YA, Meng XJ (2012) Hepatitis E virus replication requires an active ubiquitin-proteasome system. J Virol 86(10):5948–5952. https://doi.org/10.1128/JVI.07039-11
30. Suppiah S, Zhou Y, Frey TK (2011) Lack of processing of the expressed ORF1 gene product of hepatitis E virus. Virol J 8:245. https://doi.org/10.1186/1743-422X-8-245
31. Ahmad I, Holla RP, Jameel S (2011) Molecular virology of hepatitis E virus. Virus Res 161(1):47–58. https://doi.org/10.1016/j.virusres.2011.02.011
32. Xing L, Li TC, Mayazaki N, Simon MN, Wall JS, Moore M, Wang CY, Takeda N, Wakita T, Miyamura T, Cheng RH (2010) Structure of hepatitis E virion-sized particle reveals an RNA-dependent viral assembly pathway. J Biol Chem 285(43):33175–33183. https://doi.org/10.1074/jbc.M110.106336
33. Zafrullah M, Ozdener MH, Kumar R, Panda SK, Jameel S (1999) Mutational analysis of glycosylation, membrane translocation, and cell surface expression of the hepatitis E virus ORF2 protein. J Virol 73(5):4074–4082
34. Kar-Roy A, Korkaya H, Oberoi R, Lal SK, Jameel S (2004) The hepatitis E virus open reading frame 3 protein activates ERK through binding and inhibition of the MAPK phosphatase. J Biol Chem 279(27):28345–28357. https://doi.org/10.1074/jbc.M400457200

CHAPTER 6

Molecular survey of HEV infection in wild boar population in Italy

6. Molecular survey of HEV infection in wild boar population in Italy

In the present study, we analyzed the circulation of HEV in the wild boar populations hunted in 5 areas in Central Italy. Nineteen-two wild boar livers were analyzed detecting HEV RNA in 48 livers (52.2%) by Real Time RT-PCR. Twenty-one samples were subjected to nested RT-PCR to amplify a short region within ORF2. Eighteen samples from 4 areas were successfully amplified. Three samples from one area resulted negative to the amplification by conventional PCR. The strains detected the same day from the same hunter team were identical. The strains from three areas were classified as HEV-3c, HEV-3f and a group of HEV-3 equally distant from all the other subtypes and not classifiable. The Italian strains showed correlation with human and swine Italian strains previously detected in Italy. The HEV-3c strains were also correlated to the HEV-3 strain sequenced from figatelli (liver sausage) detected in France. The identical strain detected in animals hunted in the same area and infected by one identical HEV-3 strain may represent a family group while detection of different strains in animals found in the same region could be explained by the intensive hunting activities that can lead to the movement of wild boars to different area.

Received: 16 January 2018 | Revised: 30 May 2018 | Accepted: 6 June 2018

DOI: 10.1111/tbed.12948

ORIGINAL ARTICLE

WILEY Transboundary and Emerging Diseases

Molecular survey of HEV infection in wild boar population in Italy

Luca De Sabato[1,2] | Fabio Ostanello[3] | Luigi De Grossi[4] | Anita Marcario[4] | Barbara Franzetti[5] | Marina Monini[1] | Ilaria Di Bartolo[1]

[1]Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore di Sanità, Rome, Italy

[2]Department of Sciences, University Roma Tre, Rome, Italy

[3]Department of Veterinary Medical Sciences, University of Bologna, Ozzano dell'Emilia, Italy

[4]Zooprophylaxis and Research Institute of Latium and Tuscany "M. Aleandri'', Viterbo, Italy

[5]Italian National Institute for Environmental Protection and Research, Ozzano dell'Emilia, Italy

Correspondence
Ilaria Di Bartolo, Department of Veterinary Public Health and Food Safety, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy.
Email: ilaria.dibartolo@iss.it

Abstract

Hepatitis E virus (HEV) is an RNA virus causing an acute generally self-limited disease in humans. An increasing number of autochthonous cases linked to zoonotic transmission of HEV genotype 3 have been reported over the last 10 years in Europe. Pigs and wild boars are considered the main reservoirs. The principal route of transmission in Europe is food-borne, linked by direct or indirect evidence to the consumption of raw or undercooked pork products and wild boar meat. In this study, we sampled 92 wild boar (*Sus scrofa*) livers during active surveillance in five municipalities in Central Italy throughout the hunting season 2016–2017. HEV RNA was detected in 52.2% of liver sampled with prevalence ranging from 0.0% to 65.7%. HEV-positive wild boars were detected in all but one area of hunting. Phylogenetic analysis showed that strains clustered within the two subtypes HEV-3c and HEV-3f and displayed a wide range of phylogenetic diversity. Several strains were circulating in the areas investigated; animals possibly belonging to the same family group hunted by the same team were infected with a unique strain (100% nucleotide identity). As wild animals are a proven source of HEV transmission to humans and pigs, the high prevalence observed (mean 52.2%) poses a question on the risk of consuming raw or undercooked wild boar meat, and thus, this subject deserves further investigations.

KEYWORDS
genotype 3, hepatitis E virus, Italy, ngs, subtype, wild boar, zoonosis

1 | INTRODUCTION

Hepatitis E is an acute viral disease caused by hepatitis E virus (HEV) and characterized by the faecal-oral route transmission (Kamar et al., 2017). HEV is a nonenveloped single-strand RNA virus classified in the family *Hepeviridae* and the genus Orthohepevirus (Purdy et al., 2017). The genus includes the *Orthohepevirus A* species divided into seven genotypes. The genotypes HEV-1 and HEV-2, restricted to humans, circulate in endemic area (Asia and Africa) causing several outbreaks linked to the ingestion of contaminated water. In nonendemic area (industrialized countries), infections by HEV-1 and HEV-2 are related to travel in endemic area. Furthermore, in the last 10 years, an increasing number of autochthonous infections have been described linked to the zoonotic transmission of the genotypes HEV-3 and HEV-4 and are now increasingly recognized as endemic also in some developed regions. HEV-3 and HEV-4 in industrialized countries are zoonotic and linked by direct or indirect evidence to the consumption of raw pork products (mainly liver sausages) and undercooked wild boar meat (Kamar et al., 2017). The latter genotypes infect humans and several animal species among which pigs and wild boars are the main reservoirs (Ricci et al., 2017). In the recent past, novel hosts of HEV-3 and HEV-4 have been described in rabbit and yak, respectively, and novel genotypes in wild boar (HEV-5, HEV-6) and camel (HEV-7)

(Caruso, Modesto, Prato, et al., 2015; Lee et al., 2016; Smith et al., 2014; Takahashi et al., 2011; Woo et al., 2016). In Europe, HEV-3 is the most frequently detected genotype in humans, pigs and wild boars. HEV-4, mainly found in Asia, was only recently detected in Italy in pigs and in one human case (Garbuglia et al., 2013; Monne et al., 2015). The genotypes HEV-5 and HEV-6 have so far only been detected in Japanese boar (*Sus scrofa leucomystax*) (Takahashi et al., 2011). The presence of HEV-3, the most common genotype in Europe, has been extensively described in pig populations, with high seroprevalence which increases with age (up to 100%) (Pavio, Doceul, Bagdassarian, & Johne, 2017). The peak of infections in pigs is after the loss of maternal immunity: between 3 and 8 weeks of age, the virus is secreted on faeces and/or is detected in liver with prevalence ranging between 8 and 30% in weaners, 20% and 44% in growers and 8% and 73% fatteners/finishers (Pavio et al., 2017). Wild boar is also susceptible to HEV infection, displaying seroprevalences ranging between 4.9% (Caruso, Modesto, Bertolini, et al., 2015) and 57.4% (Kukielka, Rodriguez-Prieto, Vicente, & Sánchez-Vizcaíno, 2016). Among European countries, different percentages of HEV RNA detection in liver samples were also reported ranging between 3.7% (Caruso, Modesto, Bertolini, et al., 2015) and 68.2% (Adlhoch et al., 2009). Wild boar HEV-positive animals were detected in each age classes, including juveniles of 4 months of age, and animals older than 24 months (Martelli et al., 2008; Sonoda et al., 2004). In an interesting manner, a recent study described detection of HEV RNA in 89% of muscle sampled from wild boar HEV-positive in liver (Anheyer-Behmenburg et al., 2017). In Italy, the prevalence of HEV in wild boar ranges between 1.5% from faeces and 1.9% up to 33.7% from liver or bile tested. This could be a regional difference or could also be partially linked to different specimens that have been tested (faeces, bile or liver). The Italian wild boar HEV strains were sequenced in short genome regions and classified into -3c, -3e, -3f subtypes and some strains for which the subtype could not be determined, confirming the high heterogeneity of HEV in wild boars (Aprea et al., 2017; Caruso et al., 2015; Di Profio et al., 2016; Martelli et al., 2008; Martinelli et al., 2015; Mazzei et al., 2015; Montagnaro et al., 2015; Serracca et al., 2015). In this study, we investigated the occurrence of HEV in wild boars hunted in Lazio Region (Central Italy). To determine virus circulation and characterize strains detected, liver samples were tested for HEV by real-time reverse-transcription PCR (RT-qPCR) and genotyped by sequencing and phylogenetic analyses.

2 | MATERIALS AND METHODS

2.1 | Sampling

During the hunting season from October 2016 to January 2017, 92 individual liver samples were collected from wild boars (*Sus scrofa*) hunted in five municipalities (A, B, C, D and E) located in Viterbo Province (3.612 km^2, Lazio Region, Central Italy). Some geographical features of the five municipalities, about 20–50 km apart, are reported in Table 1 (ISTAT, 2018). In each municipality, hunting areas with dimensions of 25–400 hectares are assigned to specific hunting teams. In the hunting districts where the five municipalities are located, 3,374 wild boars were killed during the hunting season 2016–2017 (sex ratio 1:0.96; killing average density of 17.3 wild board/km^2 of hunting ground) (ISPRA, 2017).

2.2 | Sample preparation and nucleic acid extraction

One hundred mg of liver sample was cut with a disposable scalpel from the inner part of the organ. Samples were homogenized in 650 µl of lysis buffer (RLT) with zirconia beads, using a mechanical disruptor (Tissue Lyser, Qiagen, Milan, Italy) for three runs of 2 min at 40 oscillations s^{-1}. After centrifugation at 5,000 × *g* per 20 min, the total RNA was extracted by the RNeasy Mini kit (Qiagen, Milan, Italy), according to the manufacturer's instructions and eluted in 100 µl of nuclease-free water. Liver samples, before homogenization, were artificially contaminated with 5 µl of a suspension of murine norovirus (MuNoV, strain: MNV-IT1 Acc. no. KR349276), which was used as sample process control. The RNA of MuNoV from spiked samples was detected by Real-Time RT-PCR as previously described (Di Bartolo, Angeloni, Ponterio, Ostanello, & Ruggeri, 2015). The recovery rate was estimated by comparative cycle threshold (Ct) method (Schmittgen & Livak, 2008).

All spiked samples were positive for MuNoV, mean ± *SD* recovery rate of 12%± 8.2.

2.3 | RT-qPCR for HEV

The HEV genome was detected by quantitative Real-Time RT-PCR (RT-qPCR) as described by Jothikumar, Cromeans, Robertson, Meng, and Hill (2006) using the QuantiFast Pathogen +IC Kits (Qiagen, Milan, Italy) including the internal control (Internal Control Assay, ICA). For interpretation of results, if the observed ICA cycle threshold (Ct) value was as expected (comparable to the Ct value obtained in negative control, where only water was added as template) and the Ct value for HEV was not detectable or was ≥39, the sample was considered to be negative. Quantification of HEV genome equivalent (GE) was performed using a synthetic RNA reference standard (Di Bartolo et al., 2015). The limit of detection (LOD) was 14 GE/µl calculated using ten-fold dilution series of known amount of HEV-specific RNA molecules and defined as the lowest dilution detectable in all 10 replicates.

2.4 | Limit detection of HEV RNA

One positive homogenated liver sample (prepared in PBS) was ten-fold diluted and extracted as described above in M&M. The limit of detection, calculated as the lowest dilution with at least one positive of three triplicates, was 43.000 GE/g.

TABLE 1 HEV RNA prevalence obtained from wild boar hunted in the five examined municipalities (Zone A–E). Some geographical features are reported

		Altitude[a]						
Municipality	Area (km^2)	Min	Max	Mean	HEV RNA-positive/ examined	Prevalence (%)	95% CI	*p*
A	29.1	63	326	131	16/25	64.0	42.5–82.0	0.010
B	33.0	74	364	179	23/35	65.7	47.8–80.9	
C	105.0	220	600	400	0/5	0.0	0.0–52.2	
D	113.8	125	963	299	6/17	35.3	14.2–61.7	
E	84.2	36	213	129	3/10	30.0	6.7–65.3	
					Total 48/92	52.2	41.5–62.7	

Note. [a]Metre above sea level.

2.5 | Nested RT-PCR for HEV sequencing

The RNA (21 liver samples) was analysed by nested RT-PCR using the OneStep RT–PCR Kit (Qiagen, Milan Italy) for retro-transcription and PCR amplification and the Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific, Rodano, Italy) for the nested PCR, as previously described (Monini et al., 2015). The nested RT-PCR amplified a 348-bp region within Open Reading Frame 2 (ORF2) of HEV genome (Huang et al., 2002). The DNA amplicons were sequenced by Eurofins Genomics (Germany).

2.6 | Phylogenetic analyses

Nucleotide sequence similarity was analysed with the BLAST server (http://www.ncbi.nlm.nih.gov/genbank/index.html). A maximum likelihood (ML) phylogenetic tree was constructed with the Tamura–Nei parameter model as suggested by the MEGA 7 software model test (http://www.megasoftware.net) based on 1,000 bootstrap replications. The sequences were submitted to NCBI GenBank under accession numbers: WB02VT2016 (MG582608), WB03VT2016 (MG582609), WB17VT2016 (MG582610), WB21VT2016 (MG582611), WB27VT2016 (MG582612), WB31VT2016 (MG582613), WB35VT2016 (MG582614), WB37VT2016 (MG582615), WB39VT2016 (MG582616), WB47VT2016 (MG582617), WB52VT2016 (MG582618), WB55VT2016 (MG582619), WB57VT2016 (MG582620), WB59VT2016 (MG582621), WB61VT2016 (MG582622), WB84VT2016 (MG582623), WB89VT2016 (MG582624), WB90VT2016 (MG582625).

2.7 | Statistical analysis

Statistical analysis was performed using SPSS software (SPSS Statistics ver. 23; IBM Corp., Chicago, IL). Comparison of prevalence observed for HEV RNA-positive animals by municipality was conducted using Pearson chi-square test. The significance limit was set at $p < 0.05$. Confidence intervals were calculated by binomial (Clopper-Pearson) "exact" method based on the β distribution.

3 | RESULTS

HEV RNA was detected in liver by RT-qPCR in 52.2% (48/92; 95% C.I. 41.5–62.7) of wild boars sampled, ranging between 0.0% (area C) and 65.7% (area B; Table 1, Figure 1). A significant difference ($p < 0.05$) was observed in HEV RNA prevalence in each of the five areas (A–E) (Table 1). The median viral load was 10^7 GE/g ranging between 8.05×10^5 and 2.4×10^{10} GE/g. Twenty-one positive liver samples selected to represent at least one sample for municipalities were further analysed by conventional nested RT-PCR amplifying a 348-bp genome fragment within the ORF2 (capsid protein). Eighteen of the 21 samples were positive. However, three positive samples by RT-qPCR, all belonging to animals hunted in the area D, were not further confirmed by nested RT-PCR. Amplicons obtained were sequenced and subjected to phylogenetic analysis, including in the analyses representative reference of HEV-3 subtype strains (Smith et al., 2016), and human, swine and wild boar strains detected in both Europe and Italy available on NCBI database (https://www.ncbi.nlm.nih.gov) (Figure 2).

The wild boar strain sequences clustered according to the hunter team, displaying a high nucleotide identity 98%–100% (nt. id.) within each group of animals hunted the same day.

In the area A, two HEV strains were identified. Two identical wild boar sequences (WB02VT2016 and WB03VT2016) were assigned to HEV-3f subtype showing 89% nucleotide identity (nt. id.) with the HEV-3f prototype strain (AB369687) and up to 93% with several HEV-3f strains originated from swine, humans and wild boar (Figure 2). The two Italian wild boar strains were related to two strains detected in human cases occurring in the Netherlands and in Italy (JX645331, HM446627) and to one Italian wild boar strain (LT827027) displaying a nt. id. of 93.3%, 90.5% and 89%, respectively.

Five strain sequences from areas B and E displayed a nt. id. each of 92.5% and 92% nt. id. with the wild boar HEV-3c prototype strain (FJ705359) and were assigned to HEV-3c. HEV strains from the area B (WB17VT2016, WB21VT2016, WB27VT2016) were related (96.7% nt. id.) to both human HEV strain from the Netherlands (KR362779) and wild boar strain (KU508285) previously detected in Central Italy but in an area apparently not linked with the hunting

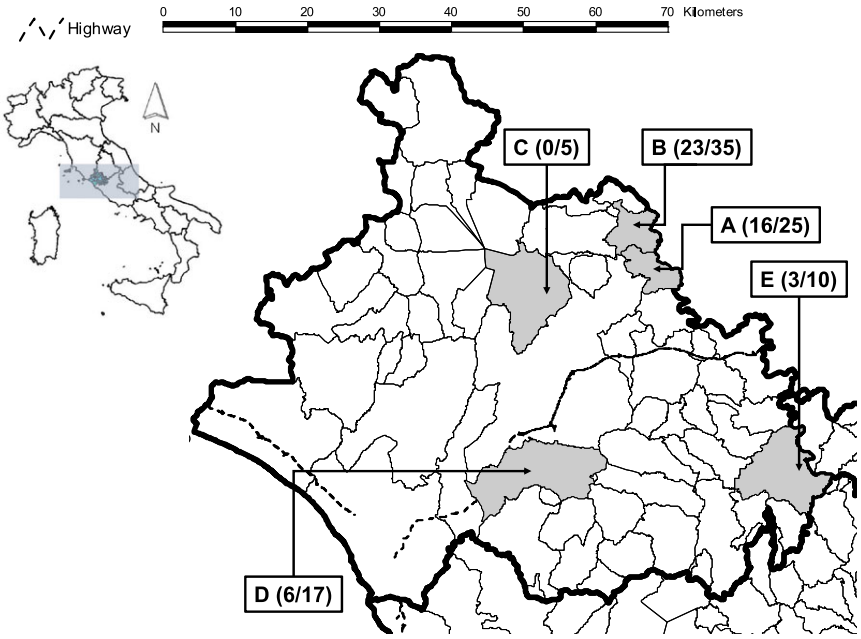


FIGURE 1 Maps of hunting area (in grey) and number of HEV-positive animals/total animals investigated per area (zone) are reported

area of the present study (Di Profio et al., 2016). Two sequence strains of HEV-3c, WB47VT2016 and WB52VT2016, from animals hunted in the area E, showed a high nt. id. (98.4%) with a strain (KF751185) detected in an acute case of hepatitis E occurring in Northern Italy (Garbuglia et al., 2015).

Four HEV strains from the area A (WB55VT2017, WB57VT2017, WB59VT2017, WB61VT2017) and seven from the area B (WB31VT2017, WB35VT2017, WB37VT2017, WB39VT2017, WB84VT2017, WB89VT2017, WB90VT2017), clustered together forming two subgroups with 95.3% nt. id., in a separate clade than the other HEV-3c (86.62% nt. id.) and out of the cluster of HEV-3hi (86.4% nt. id.). The 11 sequence strains correspond to animals hunted in three different days from three teams in areas on the edge between A and B. Sequences from animals hunted by the same team were identical (99.3-100% nt. id.) and related to two Italian wild boar strains (KX549309, LT827030) and one French human strain (KR027387) showing nt. id. 95%. Those sequences did not cluster with any reference HEV strains and were not assigned to any known subtype.

4 | DISCUSSION

In this study, the mean HEV RNA prevalence observed was 52.2%, significantly higher than in the previous studies conducted in Italy, where HEV RNA was detected in 1.9% up to 33.7% of liver or bile collected from wild boars (Martelli et al., 2008; Serracca et al., 2015). The difference can be explained for several reasons, depending on sampling strategy, the age of the examined animals, duration of storage before analyses (Schielke et al., 2009), the population density and the frequency of contact with other wild or domestic receptive species. In Italy, the area of the wild boar distribution is nearly 77% of the country (232,000 km^2) while the population size is estimated over 1.000.000 units (ISPRA, 2017). In the Lazio region where animals investigated in this study were hunted, the wild boar population is distributed over 60% of the territory and the number of animals hunted per year is probably lower than the annual growth rate (ARSIAL, Osservatorio Faunistico Regionale, Lazio, DAFNE, Università degli Studi della, & Tuscia, 2014). The high density and the contact with other animal species are considered risk factors for several infections. In the studied areas, during the hunting season 2016-2017, an average of 17.3 wild boars/km^2 was hunted (ISPRA, 2017). This high density of hunted wild boar could partially justify the observed high HEV prevalence, although this hypothesis in the absence of the exact density value has to be proved yet. No sequence data are available on area D because the three samples positive by RT-qPCR were not further confirmed by nested RT-PCR. We suppose that this is due heterogeneity of strain sequences that we will investigate in the future.

According to the national and regional regulations, hunting on grounds is allowed from October to January. In the present study, animals were hunted during four months in five small areas, 20-50 km apart, separated by geographical barriers (Table 1). Strains detected from animals hunted on the same day by the same team showed 100% nt. id. The other HEV strains detected were shown to be different both among the different hunting areas and within the same hunting area. This result is interesting because we are able to conclude that more strains were circulating but that animals belonging to the same family group shared one unique strain (100% nt. id.). Wild boars hunted on the same day by the same hunting team may belong to the same family group, as except for the old males, wild boar live in groups consisting of interrelated females and their litters

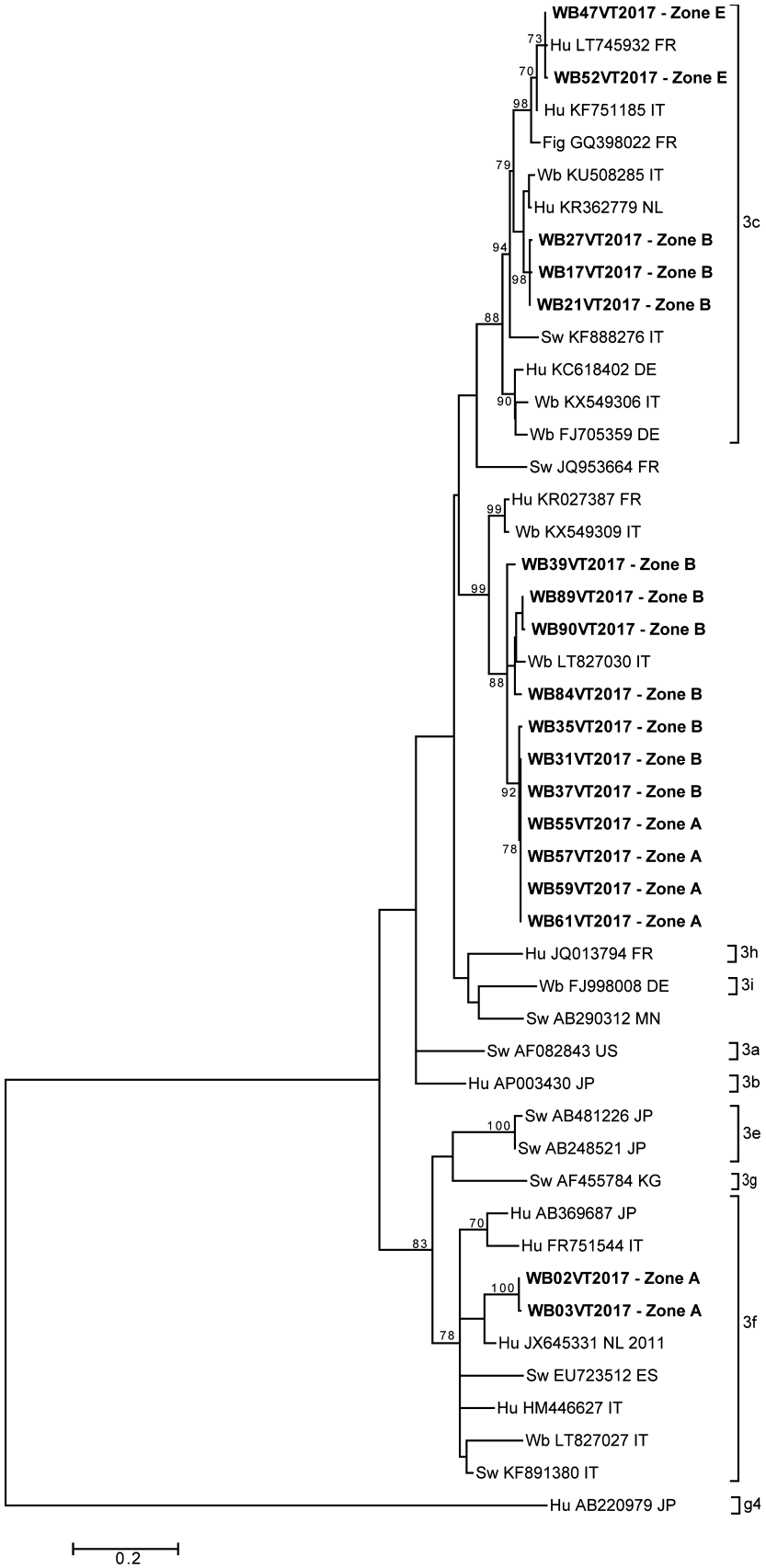


FIGURE 2 Phylogenetic tree based on the 302-bp sequences of the ORF2 fragment. Representative porcine, human and wild boar strains are included. Each entry includes host (Fig: figatelli; Hu: human, Sw: swine, Wb: wild boar), accession number and countries origin of strains. Strains detected in this study are in bold. Bootstrap values >70 are indicated

(Briedermann, 1986; Kaminski, Brandt, Baubet, & Baudoin, 2005). However, some studies reported aggregations of unrelated adult females (with their litters) in family group, mainly due to intensive hunting activity breaking down the structure of family groups (Brün & Keuling, 2008; Gabor, Hellgren, Van Den Bussche, & Silvy, 1999; Iacolina, 2009). Indeed, we detected the same strains circulating in wild boars hunted along the border between A and B areas, where the probability of contact between animals can be considered high (i.e., WB35VT2017, WB61VT2017, WB31VT2017, WB37VT2017, WB55VT2017, WB57VT2017, WB59VT2017). However, we have also detected different HEV strains circulating in the same area (e.g., WB27VT2017 vs WB31VT2017). As explained above, the intensive hunting can determine movement of animals escaping from one area to other, joining to new family group. This could increase contact among animals explaining different HEV strains detected within the same area. We may assume that animals (from areas A, B, C and from areas E and D) belong to different metapopulations with limited contact with other groups or subpopulations. Supporting this hypothesis, different HEV strains were also observed within the same area. Wild boars are sedentary and move very little <10–12 km throughout the day (Morelle et al., 2015). Furthermore, the geographical distances that separate the different sampling areas (e.g., areas A, B, C and areas E and D¸ 30 km between A and E, ¸20 km between D and E and¸ 15 km between C and D) are relatively far apart.

The age of animals was not available, but if we consider that the peak of births is in spring, we expect that during the hunting season (October–January) wild boars are older than 6 months. Wild boars of this age are those usually intended for human consumption. This result confirms previous findings that adult pigs showed a lower probability of infection (Di Bartolo et al., 2008) while wild boars can be infected at different ages (Martelli et al., 2008). This can be linked to the chronic infection described in wild boar (Schlosser, Vina-Rodriguez, Fast, Groschup, & Eiden, 2015) or continuous re-infection due to incomplete or short-lasting immunity (Anheyer-Behmenburg et al., 2017).

In this study, we observed a median viral load of 10^7 GE/gr. This value is comparable to previous studies (Anheyer-Behmenburg et al., 2017; Kamar et al., 2017). In the absence of in vitro cultivation, detection of HEV RNA does not confirm the viability of the virus. However, the observed viral load in liver, that is the main site of virus replication, deserves attention as liver is also used to produce regional food specialties such as liver sausages that could be consumed raw.

This study confirmed a wide heterogeneity of sequenced wild boar HEV strains that belonged to at least two subtypes HEV-3c and HEV-3f. The sequence analyses revealed that HEV-3c is frequent, as observed in other studies conducted in both Italy and Europe (Anheyer-Behmenburg et al., 2017; Aprea et al., 2017; Di Profio et al., 2016; Dorn-In et al., 2017; Rutjes et al., 2009, 2010; Schielke et al., 2009; Serracca et al., 2015; Thiry, Mauroy, et al., 2017; Vina-Rodriguez et al., 2015). In Italy, HEV-3c is less frequent detected in pigs, where the main subtypes are HEV-3f and HEV3-e. Eleven HEV strains detected from animals hunted in the border of area A and B shows nucleotide identity <86.5% with reference sequences HEV-3c and HEV-3-hi (Figure 2), suggesting a possible local evolution but not allowing a definitive assignment to the subtypes known so far.

Furthermore, HEV strains sequenced in this study displayed higher nucleotide identity with human and wild boar strains than with HEV strains detected in pigs.

We observed a high nucleotide identity with a human strain detected in Italy linked to consumption of figatelli (raw pork liver sausages) (Garbuglia et al., 2015). Human HEV infection after ingestion of uncooked liver and meat of wild boars was reported in Japan and Spain, respectively (Li et al., 2005; Rivero-Juarez et al., 2017). In Italy, one human case (Giordani, Fabris, Brunetti, Goblirsch, & Romanò, 2013) was supposed to be linked to wild boar meat consumption because the patient had never travelled outside Italy and declared to have consumed wild boar meat. In the same area, wild boar HEV strains related to those detected in the human case were also reported (Mazzei et al., 2015). Wild boars might represent a source of autochthonous HEV transmission to humans in those regions where consumption of wild boar meat is common or where there is frequent contact between pigs and wild boars. Indeed, the transmission of HEV between domestic and wild swine has been clearly demonstrated (Thiry, Rose, et al., 2017). Moreover, wild boars intended for human consumption are mainly captured by hunting and game meat follows a food chain different from pigs, where rules for food safety could be less strict. Furthermore, the exposure to wild boar carcasses could be a relevant source of risks for hunters (Schielke et al., 2015). Our findings suggest that wild boar consumption and circulation of HEV in sylvatic populations deserve further investigation and special attention by wildlife managers, veterinarian and hunters.

ACKNOWLEDGEMENTS

We would like to acknowledge Edoardo Vignolo for his valuable technical support and the ATCVT1 and ATCVT2 for their support in gathering data.

CONFLICT OF INTEREST

No conflict of interest to declare.

ORCID

Luca De Sabato http://orcid.org/0000-0001-9995-8195
Fabio Ostanello http://orcid.org/0000-0003-1122-8628
Ilaria Di Bartolo http://orcid.org/0000-0001-5674-1332

REFERENCES

Adlhoch, C., Wolf, A., Meisel, H., Kaiser, M., Ellerbrok, H., & Pauli, G. (2009). High HEV presence in four different wild boar populations in

East and West Germany. *Veterinary Microbiology*, *139*(3–4), 270–278. https://doi.org/10.1016/j.vetmic.2009.06.032

Anheyer-Behmenburg, H. E., Szabo, K., Schotte, U., Binder, A., Klein, G., & Johne, R. (2017). Hepatitis E virus in wild boars and spillover infection in red and roe deer, Germany, 2013-2015. *Emerging Infectious Diseases*, *23*(1), 130–133. https://doi.org/10.3201/eid2301.161169

Aprea, G., Amoroso, M. G., Di Bartolo, I., D'Alessio, N., Di Sabatino, D., Boni, A., … Fusco, G. (2017). Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy. *Transboundary and Emerging Diseases*, https://doi.org/10.1111/tbed.12661

ARSIAL, Osservatorio Faunistico Regionale, Lazio, R., & DAFNE, Università degli Studi della Tuscia (2014). Indagine sullo status del cinghiale (*Sus scrofa* L.) nel Lazio e sulle interazioni con le attività antropiche: base conoscitiva per l'assestamento faunistico venatorio della specie nel Lazio – Relazione preliminare – luglio 2014.

Briedermann, L. (1986). Schwarzwild (Wild boar). VEB Deutscher, Landwirtschaftsverlag Berlin.

Brün, J., & Keuling, O. (2008). Socio-spatial behaviour and genetics of wild boar in North-Eastern Germany. Abstracts of the 7th International Symposium on Wild Boar (*Sus scrofa*) and Sub-order Suiformes.

Caruso, C., Modesto, P., Bertolini, S., Peletto, S., Acutis, P. L., Dondo, A., … Masoero, L. (2015). Serological and virological survey of hepatitis E virus in wild boar populations in northwestern Italy: Detection of HEV subtypes 3e and 3f. *Archives of Virology*, *160*(1), 153–160. https://doi.org/10.1007/s00705-014-2246-5

Caruso, C., Modesto, P., Prato, R., Scaglione, F. E., De Marco, L., Bollo, E., … Peletto, S. (2015). Hepatitis E virus: First description in a pet house rabbit. A new transmission route for human? *Transboundary and Emerging Diseases*, *62*(3), 229–232. https://doi.org/10.1111/tbed.12348

Di Bartolo, I., Angeloni, G., Ponterio, E., Ostanello, F., & Ruggeri, F. M. (2015). Detection of hepatitis E virus in pork liver sausages. *International Journal of Food Microbiology*, *193*, 29–33. https://doi.org/10.1016/j.ijfoodmicro.2014.10.005

Di Bartolo, I., Martelli, F., Inglese, N., Pourshaban, M., Caprioli, A., Ostanello, F., & Ruggeri, F. M. (2008). Widespread diffusion of genotype 3 hepatitis E virus among farming swine in Northern Italy. *Veterinary Microbiology*, *132*(1–2), 47–55. https://doi.org/10.1016/j.vetmic.2008.04.028

Di Profio, F., Melegari, I., Sarchese, V., Robetto, S., Marruchella, G., Bona, M. C., … Di Martino, B. (2016). Detection and genetic characterization of hepatitis E virus (HEV) genotype 3 subtype c in wild boars in Italy. *Archives of Virology*, *161*(10), 2829–2834. https://doi.org/10.1007/s00705-016-2964-y

Dorn-In, S., Schwaiger, K., Twarużek, M., Grajewski, J., Gottschalk, C., & Gareis, M. (2017). Hepatitis E virus in wild boar in Northwest Poland: Sensitivity of methods of detection. *Foodborne Pathogens and Disease*, *14*(2), 103–108. https://doi.org/10.1089/fpd.2016.2194

Gabor, T. M., Hellgren, E. C., Van Den Bussche, R. A., & Silvy, N. J. (1999). Demography, sociospatial behaviour and genetics of feral pigs (*Sus scrofa*) in a semi-arid environment. *Journal of Zoology*, *247*, 311–322.

Garbuglia, A. R., Alessandrini, A. I., Pavio, N., Tessé, S., Grignolo, S., Viscoli, C., … Capobianchi, M. R. (2015). Male patient with acute hepatitis E in Genoa, Italy: Figatelli (pork liver sausage) as probable source of the infection. *Clinical Microbiology and Infection*, *21*(1), e4–e6. https://doi.org/10.1016/j.cmi.2014.07.007

Garbuglia, A. R., Scognamiglio, P., Petrosillo, N., Mastroianni, C. M., Sordillo, P., Gentile, D., … Capobianchi, M. R. (2013). Hepatitis E virus genotype 4 outbreak, Italy, 2011. *Emerging Infectious Diseases*, *19*(1), 110–114. https://doi.org/10.3201/eid1901.120983

Giordani, M. T., Fabris, P., Brunetti, E., Goblirsch, S., & Romanò, L. (2013). Hepatitis E and lymphocytic leukemia in Man, Italy. *Emerging Infectious Diseases*, *19*(12), 2054–2056. https://doi.org/10.3201/eid1912.130521

Huang, F. F., Haqshenas, G., Guenette, D. K., Halbur, P. G., Schommer, S. K., Pierson, F. W., … Meng, X. J. (2002). Detection by reverse transcription-PCR and genetic characterization of field isolates of swine hepatitis E virus from pigs in different geographic regions of the United States. *Journal of Clinical Microbiology*, *40*(4), 1326–1332. https://doi.org/10.1128/JCM.40.4.1326-1332.2002

Iacolina, L. (2009). Nonkin Associations in Wild Boar Social Units. *Journal of Mammalogy*, *90*(3), 666–674.

ISPRA (2017). Italian National Institute for Environmental Protection and Research. Retrieved from http://www.isprambiente.gov.it/it/banche-dati.

ISTAT (2018). Italian National Institute of Statistics. Retrieved from http://www.istat.it/it/archivio/156224

Jothikumar, N., Cromeans, T. L., Robertson, B. H., Meng, X. J., & Hill, V. R. (2006). A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. *Journal of Virological Methods*, *131*(1), 65–71. https://doi.org/10.1016/j.jviromet.2005.07.004

Kamar, N., Izopet, J., Pavio, N., Aggarwal, R., Labrique, A., Wedemeyer, H., & Dalton, H. R. (2017). Hepatitis E virus infection. *Nature Reviews Disease Primers*, *3*, 17086. https://doi.org/10.1038/nrdp.2017.86

Kaminski, G., Brandt, S., Baubet, E., & Baudoin, C. (2005). Life-history patterns in female wild boars (*Sus scrofa*): Mother–daughter post-weaning associations. *Canadian Journal of Zoology*, *93*(3), 474–480.

Kukielka, D., Rodriguez-Prieto, V., Vicente, J., & Sánchez-Vizcaíno, J. M. (2016). Constant hepatitis E virus (HEV) circulation in wild boar and red deer in spain: An increasing concern source of HEV zoonotic transmission. *Transboundary and Emerging Diseases*, *63*(5), e360–e368. https://doi.org/10.1111/tbed.12311

Lee, G.-H., Tan, B.-H., Chi-Yuan Teo, E., Lim, S.-G., Dan, Y.-Y., Wee, A., … Teo, C.-G. (2016). Chronic infection with camelid hepatitis E virus in a liver transplant recipient who regularly consumes camel meat and milk. *Gastroenterology*, *150*(2), 355–357. e353. https://doi.org/10.1053/j.gastro.2015.10.048

Li, T.-C., Chijiwa, K., Sera, N., Ishibashi, T., Etoh, Y., Shinohara, Y., … Miyamura, T. (2005). Hepatitis E virus transmission from wild boar meat. *Emerging Infectious Diseases*, *11*(12), 1958–1960. https://doi.org/10.3201/eid1112.051041

Martelli, F., Caprioli, A., Zengarini, M., Marata, A., Fiegna, C., Di Bartolo, I., … Ostanello, F. (2008). Detection of hepatitis E virus (HEV) in a demographic managed wild boar (*Sus scrofa* scrofa) population in Italy. *Veterinary Microbiology*, *126*(1–3), 74–81. https://doi.org/10.1016/j.vetmic.2007.07.004

Martinelli, N., Pavoni, E., Filogari, D., Ferrari, N., Chiari, M., Canelli, E., & Lombardi, G. (2015). Hepatitis E virus in wild boar in the central northern part of Italy. *Transboundary and Emerging Diseases*, *62*(2), 217–222. https://doi.org/10.1111/tbed.12118

Mazzei, M., Nardini, R., Verin, R., Forzan, M., Poli, A., & Tolari, F. (2015). Serologic and molecular survey for hepatitis E virus in wild boar (*Sus scrofa*) in Central Italy. *New Microbes and New Infections*, *7*, 41–47. https://doi.org/10.1016/j.nmni.2015.05.008

Monini, M., Di Bartolo, I., Ianiro, G., Angeloni, G., Magistrali, C. F., Ostanello, F., & Ruggeri, F. M. (2015). Detection and molecular characterization of zoonotic viruses in swine fecal samples in Italian pig herds. *Archives of Virology*, *160*(10), 2547–2556. https://doi.org/10.1007/s00705-015-2538-4

Monne, I., Ceglie, L., di Martino, G., Natale, A., Zamprogna, S., Morreale, A., … Bonfanti, L. (2015). Hepatitis E virus genotype 4 in a pig farm, Italy, 2013. *Epidemiology and Infection*, *143*(3), 529–533. https://doi.org/10.1017/S0950268814001150

Montagnaro, S., De Martinis, C., Sasso, S., Ciarcia, R., Damiano, S., Auletta, L., … Pagnini, U. (2015). Viral and antibody prevalence of hepatitis E in European Wild Boars (*Sus scrofa*) and Hunters at Zoonotic Risk in the Latium Region. *Journal of Comparative Pathology*, *153*(1), 1–8. https://doi.org/10.1016/j.jcpa.2015.04.006

Morelle, K., Podgórski, T., Prévot, C., Keuling, O., Lehaire, F., & Lejeune, P. (2015). Towards understanding wild boar *Sus scrofa* movement: A synthetic movement ecology approach. *Mammal Review*, *45*, 15–29.

Pavio, N., Doceul, V., Bagdassarian, E., & Johne, R. (2017). Recent knowledge on hepatitis E virus in Suidae reservoirs and transmission routes to human. *Veterinary Research*, *48*(1), 78. https://doi.org/10.1186/s13567-017-0483-9

Purdy, M. A., Harrison, T. J., Jameel, S., Meng, X.-J., Okamoto, H., … Van der Poel, W. H. M. (2017). ICTV virus taxonomy profile: Hepeviridae. *The Journal of General Virology*, *98*(11), 2645–2646. https://doi.org/10.1099/jgv.0.000940

Ricci, A., Allende, A., Bolton, D., Chemaly, M., Davies, R., Fernandez Escamez, P. S., … Robertson, L. (2017). Public health risks associated with hepatitis E virus (HEV) as a food-borne pathogen. *EFSA Journal*, *15*(7), https://doi.org/Unsp 488610.2903/J.Efsa. 2017.4886

Rivero-Juarez, A., Frias, M., Martinez-Peinado, A., Risalde, M. A., Rodriguez-Cano, D., Camacho, A., … Rivero, A. (2017). Familial hepatitis E outbreak linked to wild boar meat consumption. *Zoonoses and Public Health*, *64*(7), 561–565. https://doi.org/10.1111/zph.12343

Rutjes, S. A., Lodder, W. J., Lodder-Verschoor, F., van den Berg, H. H. J. L., Vennema, H., Duizer, E., … de Roda Husman, A. M. (2009). Sources of hepatitis E virus genotype 3 in The Netherlands. *Emerging Infectious Diseases*, *15*(3), 381–387. https://doi.org/10.3201/eid1503.071472

Rutjes, S. A., Lodder-Verschoor, F., Lodder, W. J., van der Giessen, J., Reesink, H., Bouwknegt, M., & de Roda Husman, A. M. (2010). Seroprevalence and molecular detection of hepatitis E virus in wild boar and red deer in The Netherlands. *Journal of Virological Methods*, *168*(1–2), 197–206. https://doi.org/10.1016/j.jviromet.2010.05.014

Schielke, A., Ibrahim, V., Czogiel, I., Faber, M., Schrader, C., Dremsek, P., … Johne, R. (2015). Hepatitis E virus antibody prevalence in hunters from a district in Central Germany, 2013: A cross-sectional study providing evidence for the benefit of protective gloves during disembowelling of wild boars. *BMC Infectious Diseases*, *15*, 440. https://doi.org/10.1186/s12879-015-1199-y

Schielke, A., Sachs, K., Lierz, M., Appel, B., Jansen, A., & Johne, R. (2009). Detection of hepatitis E virus in wild boars of rural and urban regions in Germany and whole genome characterization of an endemic strain. *Virology Journal*, *6*, 58. https://doi.org/10.1186/1743-422X-6-58

Schlosser, J., Vina-Rodriguez, A., Fast, C., Groschup, M. H., & Eiden, M. (2015). Chronically infected wild boar can transmit genotype 3 hepatitis E virus to domestic pigs. *Veterinary Microbiology*, *180*(1–2), 15–21. https://doi.org/10.1016/j.vetmic.2015.08.022

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols*, *3*(6), 1101–1108.

Serracca, L., Battistini, R., Rossini, I., Mignone, W., Peletto, S., Boin, C., … Ercolini, C. (2015). Molecular investigation on the presence of hepatitis E virus (HEV) in Wild Game in North-Western Italy. *Food and Environmental Virology*, *7*(3), 206–212. https://doi.org/10.1007/s12560-015-9201-9

Smith, D. B., Simmonds, P., Jameel, S., Emerson, S. U., Harrison, T. J., … Meng, X.-J. (2014). Consensus proposals for classification of the family Hepeviridae. *The Journal of General Virology*, *95*(Pt 10), 2223–2232. https://doi.org/10.1099/vir.0.068429-0

Smith, D. B., Van der Poel, W. H. M., Meng, X.-J., Koenig, M., Harrison, T. J., Oliveira-Filho, E. F., … Izopet, J. (2016). Proposed reference sequences for hepatitis E virus subtypes. *The Journal of General Virology*, *97*(3), 537–542. https://doi.org/10.1099/jgv.0.000393

Sonoda, H., Abe, M., Sugimoto, T., Sato, Y., Bando, M., Fukui, E., … Okamoto, H. (2004). Prevalence of hepatitis E virus (HEV) Infection in wild boars and deer and genetic identification of a genotype 3 HEV from a boar in Japan. *Journal of Clinical Microbiology*, *42*(11), 5371–5374. https://doi.org/10.1128/JCM.42.11.5371-5374.2004

Takahashi, M., Nishizawa, T., Sato, H., Sato, Y., Jirintai, S., Nagashima, S., & Okamoto, H. (2011). Analysis of the full-length genome of a hepatitis E virus isolate obtained from a wild boar in Japan that is classifiable into a novel genotype. *The Journal of General Virology*, *92*(Pt 4), 902–908. https://doi.org/10.1099/vir.0.029470-0

Thiry, D., Mauroy, A., Saegerman, C., Licoppe, A., Fett, T., Thomas, I., … Linden, A. (2017). Belgian wildlife as potential zoonotic reservoir of hepatitis E virus. *Transboundary and Emerging Diseases*, *64*(3), 764–773. https://doi.org/10.1111/tbed.12435

Thiry, D., Rose, N., Mauroy, A., Paboeuf, F., Dams, L., Roels, S., … Thiry, E. (2017). Susceptibility of pigs to zoonotic hepatitis E virus genotype 3 isolated from a wild boar. *Transboundary and Emerging Diseases*, *64*(5), 1589–1597. https://doi.org/10.1111/tbed.12550

Vina-Rodriguez, A., Schlosser, J., Becher, D., Kaden, V., Groschup, M., & Eiden, M. (2015). Hepatitis E virus genotype 3 diversity: Phylogenetic analysis and presence of subtype 3b in wild boar in Europe. *Viruses*, *7*(5), 2704–2726. https://doi.org/10.3390/v7052704

Woo, P. C. Y., Lau, S. K. P., Teng, J. L. L., Cao, K.-Y., Wernery, U., Schountz, T., … Yuen, K.-Y. (2016). New hepatitis E virus genotype in bactrian camels, Xinjiang, China, 2013. *Emerging Infectious Diseases*, *22*(12), 2219–2221. https://doi.org/10.3201/eid2212.160979

How to cite this article: De Sabato L, Ostanello F, De Grossi L, et al. Molecular survey of HEV infection in wild boar population in Italy. *Transbound Emerg Dis*. 2018;00:1–8. https://doi.org/10.1111/tbed.12948

CHAPTER 7

Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy

7. Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy

Bats are the natural reservoir of several Coronaviruses species. During the last decade, two Beta-CoVs, MERS-CoV and SARS-CoV, have caused thousands of cases and deaths worldwide causing severe respiratory diseases in human. After the SARS outbreak and the discovering of Rhinopolus bats as the natural reservoirs of SARS-CoVs strains, the surveillance of CoVs in bats has exponentially increased. During the MERS outbreaks, camels have been described as zoonotic reservoir of viral strains infecting human. Strains detected in camels and human cases show a high genetic correlation (>98% nucleotide identity). Phylogenetic studies on bat CoV strains detected in China and Africa showed the evolutionary correlation of bat CoVs to human and camel strains. Despite the classification in the same CoV species, bat strains, recognized as MERS-CoV like strains, showed <90% nucleotide identity with human and camel strains. In Europe and in Italy, some studies reported the detection of several species of CoVs belonging to the two Alpha and Beta-CoVs genera in bats. A first classification criteria was based on sequencing and phylogenetic analysis of the short conserved fragment within the RdRp region (ORF1ab). However, the sequence analysis of this region is often hampered by the high heterogeneity and genetic diversity of CoV species. To help with the classification of CoV strains, the ICTV has proposed a classification criteria based on the analyses of the most conserved regions inside the ORF1ab. Furthermore, the complete genome sequences and the analysis of accessory proteins help to understand the evolutionary correlation of CoVs and the protein structure prediction of the Spike protein may help to understand the binding potential to human receptor. In the present study, two strains detected from *Pipistrellus kuhlii* and *Hypsugo savii* were fully sequenced. The phylogenetic analysis showed two Beta-CoVs related to MERS-CoV strains and classified in the same species. The applied NGS method and the phylogenetic analysis enables comprehensive characterization of CoVs. The surveillance on bat populations can help to understand better the CoV circulating species and their evolution.

Moreno et al. Virology Journal (2017) 14:239
DOI 10.1186/s12985-017-0907-1

Virology Journal

RESEARCH Open Access

Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy

Ana Moreno[1*], Davide Lelli[1], Luca de Sabato[2,3], Guendalina Zaccaria[2], Arianna Boni[2], Enrica Sozzi[1], Alice Prosperi[1], Antonio Lavazza[1], Eleonora Cella[4], Maria Rita Castrucci[2], Massimo Cicozzi[4] and Gabriele Vaccari[2]

Abstract

Background: Middle East respiratory syndrome coronavirus (MERS-CoV), which belongs to beta group of coronavirus, can infect multiple host species and causes severe diseases in humans. Multiple surveillance and phylogenetic studies suggest a bat origin. In this study, we describe the detection and full genome characterization of two CoVs closely related to MERS-CoV from two Italian bats, *Pipistrellus kuhlii* and Hypsugo savii.

Methods: Pool of viscera were tested by a pan-coronavirus RT-PCR. Virus isolation was attempted by inoculation in different cell lines. Full genome sequencing was performed using the Ion Torrent platform and phylogenetic trees were performed using IQtree software. Similarity plots of CoV clade c genomes were generated by using SSE v1.2. The three dimensional macromolecular structure (3DMMS) of the receptor binding domain (RBD) in the S protein was predicted by sequence-homology method using the protein data bank (PDB).

Results: Both samples resulted positive to the pan-coronavirus RT-PCR (IT-batCoVs) and their genome organization showed identical pattern of MERS CoV. Phylogenetic analysis showed a monophyletic group placed in the Beta2c clade formed by MERS-CoV sequences originating from humans and camels and bat-related sequences from Africa, Italy and China. The comparison of the secondary and 3DMMS of the RBD of IT-batCoVs with MERS, HKU4 and HKU5 bat sequences showed two aa deletions located in a region corresponding to the external subdomain of MERS-RBD in IT-batCoV and HKU5 RBDs.

Conclusions: This study reported two beta CoVs closely related to MERS that were obtained from two bats belonging to two commonly recorded species in Italy (*P. kuhlii* and H. savii). The analysis of the RBD showed similar structure in IT-batCoVs and HKU5 respect to HKU4 sequences. Since the RBD domain of HKU4 but not HKU5 can bind to the human DPP4 receptor for MERS-CoV, it is possible to suggest also for IT-batCoVs the absence of DPP4-binding potential. More surveillance studies are needed to better investigate the potential intermediate hosts that may play a role in the interspecies transmission of known and currently unknown coronaviruses with particular attention to the S protein and the receptor specificity and binding affinity.

Keywords: MERS-like Beta-CoV viruses, Full genome sequencing, Bats, Italy, Phylogenetic and molecular analyses

* Correspondence: anamaria.morenomartin@izsler.it
[1]Department of Virology, Istituto Zooprofilattico Sperimentale Lombardia ed Emilia Romagna, Via Antonio Bianchi 9, 25124 Brescia, Italy
Full list of author information is available at the end of the article

Background

Since the early 70s, a variety of pathological conditions in domestic and wild animals have been attributed to coronavirus (CoV) infections. Currently, six different CoV strains are known to infect humans [1]. Two of these belong to the beta CoV genus, severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), and they cause severe respiratory diseases with case fatality rates of 9% and 35%, respectively [2]. The reservoir of these viruses is usually animal with occasional spillover into humans, possibly through an intermediate host species. Apart from animal to human transmission, human-to-human transmission of SARS-CoV and MERS-CoV occurs mainly through nosocomial transmission [3]. Bats, with their extensive geographical distribution and flight capability, have been documented as natural hosts of a large number of diverse viruses, such as lyssaviruses, paramyxoviruses and filoviruses. Moreover, the genetic diversity of CoVs in bats exceeds that known for other hosts, which is compatible with bats being the major reservoir of mammalian CoVs [4].

The evolutionary origin of SARS-CoV, which was first detected in 2002, involved bat hosts, possibly with civets as intermediate host and the source of human infection [4]. The origin of MERS-CoV is not well known, but more recent studies point to camels as possible reservoirs or intermediate hosts. Bats have also been suspected as the evolutionary source of MERS-CoV due to the genetic similarities between beta CoVs found in bats and the MERS-CoV in humans [5, 6].

The receptor binding of CoV is mediated by the Spike protein (S), which is further cleaved into S1 and S2 subunits that are involved in engaging receptors and mediating membrane fusion, respectively. The peptidase recognized by MERS-CoV was identified as dipeptidyl peptidase 4 (DPP4 or CD26) [7]. The S1 domain responsible for the recognition of DPP4 receptor is located in a C-terminal 240-residue receptor binding domain and is composed of a core and an external subdomain. This external subdomain, also designated as the receptor binding motif, engages the receptor. Investigation of the DPP4-binding potential of bat CoVs is essential to better understanding the biology of these viruses, the eventual role in the evolutionary pathway of MERS-CoV and their potential threat to human health. Although high sequence identity in S protein was observed between BatCoVs HKU4/HKU5 and MERS-CoV, it was recently demonstrated that only the RBD of HKU4 was able to bind the human receptor DPP4 [8]. Even if it is less adapted than MERS-RBD and shows lower affinity for receptor binding, the ability of HKU4 to bind human DPP4 indicates its potential for adaptation to infect humans.

On the other hand, other authors [9] reported that MERS-RBD interacts efficiently with Jamaican fruit bat DPP4 receptor and MERS-CoV replicates efficiently in Jamaican fruit bat cells, suggesting that there is no restriction at the receptor or cellular level for MERS-CoV.

A variety of closely MERS-related CoV sequences have been obtained from numerous bat species in different continents. A fragment of a CoV showing 100% identity to HCoV-EMC/2012 cloned from the index MERS case was found in a faecal sample from an Egyptian tomb bat (*Taphozous perforatus*) in Bisha, South Arabia [5]. Partial genome sequences from viruses closely related to MERS-CoV have also been detected in bats from Africa, America and Europe [10–13]. CoVs originated from bats in Africa [6, 14], and in China [15], they were fully sequenced and identified as highly related to MERS-CoV. Despite all these reports, only three regarded the complete genome [6, 14, 15]. The other ones were based on short genomic sequences of a conserved fragment codifying the RNA-dependent RNA polymerase (RdRp) gene, which is less informative and unsuitable for solid phylogenetic hypothesis. In this study, we describe the detection and full genome characterization of two CoVs closely related to MERS-CoV from Italian bats of different species, *Pipistrellus kuhlii* and *Hypsugo savii*.

Methods

Sampling

Fresh carcasses of spontaneously dead bats were obtained from a wildlife rehabilitation centre in the context of a virological survey implemented in Northern Italy since 2010. The identification of bat species was made according to morphologic characteristics reported in the illustrated identification key to the bats of Europe [16].

Pan-coronavirus RT-PCR

Pools of viscera (lung, heart, spleen and liver) and intestine were homogenized in minimal essential medium (MEM, 1 g/10 ml) containing antibiotics and clarified by centrifugation at 3000×g for 15 min. Viral RNA was extracted from 100 μl of sample using the NucleoMag 96 Virus kit (Macherey-Nagel, Düren, Germany). The RNA was eluted in 100 μl of MV6 elution buffer and stored at −80 °C. CoV screening was performed by a pan-coronavirus one-step RT-PCR method based on degenerate primers that amplified a fragment (180 bp) of the RNA-dependent RNA polymerase (RdRp) gene [12].

Virus isolation attempts

Virus isolation was attempted by inoculation with tissue samples of different cell lines such as VERO cells (African green monkey kidney), MARC-145 (foetal monkey kidney), HRT-18 (human colorectal adenocarcinoma), FRhK 4 (foetal rhesus kidney), LLC-Mk2 (rhesus monkey

kidney) and TB1 LU (lung, Mexican free-tailed bat, "*Tadarida brasiliensis* mexicana"). Confluent monolayers of cell lines were inoculated with samples, incubated at 37 °C with 5% CO_2 and observed daily for seven days for the development of cytopathic effects.

For CoV isolation, cell cultures were used with growth media (Eagle's minimum essential medium (EMEM)) supplemented with 10% fetal bovine plus antibiotics (100 units/ml of penicillin and 100 μg/ml of streptomycin), 0.3% tryptose phosphate broth (Sigma, USA), 0.02% yeast extract (Sigma, USA) and 10 μg/ml trypsin. Twenty-four well tissue culture plates were inoculated with 0.2 ml per well of the clarified pathological material. After adsorption for 1 h at 37 °C, maintenance medium EMEM supplemented with 1% fetal bovine serum and antibiotics (0.8 ml per well) was added without removing the viral inoculum, and the cultures were incubated at 37 °C.

Whole-genome sequencing

Libraries were prepared following sequence independent single primer amplification (SISPA) with several variations as described by Djikeng et al. [17].

Nine microlitres of extracted RNA was used for reverse transcription reaction using a combination of random (FR26RV-N) and poly T (FR40RV-T) primers tagged with the sequence 5'-GCC GGA GCT CTG CAG ATA TC-3′, using SuperScript III Reverse Transcriptase (Invitrogen, Monza, Italy) following the manufacturer's instructions.

The second strand of cDNA was synthesized by DNA Polymerase I Large (Klenow) Fragment (Promega, Milan, Italy) using 20 μl of cDNA. Twenty microlitres of Klenow product was amplified by the Expand High Fidelity PCR System (Sigma Aldrich S.R.L., Milan, Italy) using FR20RV-T primer complementary to the sequence tag. Five microlitres of the PCR product was analysed on a 1% agarose gel. The PCR amplicons were purified using the QIAquick PCR Purification Kit (Qiagen, Milan, Italy) following the manufacturer's instructions and eluted in 40 μl of nuclease-free water.

The purified DNA was quantified in the Qubit 2.0 Fluorometer (Invitrogen) using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Rodano, MI, Italy) following the manufacturer's instructions. Five hundred nanograms of purified DNA were digested with EcoRV enzyme (New England BioLabs, Pero, MI, Italy) to remove the tag sequences. Digested DNA was cleaned up adding a 1.8× volume of Agencourt AMPure XP beads (Beckman, Milan, Italy). DNA was quantified by Quibit 2.0 Fluorometer and the library prepared by Ion Xpress Plus gDNA Fragment Library Preparation (Thermo Fisher Scientific) following the standard protocol for 100 ng of DNA. Emulsion PCR was performed using the Ion PGM Template OT2 200 Kit and the sequencing run performed ac-cording to the instructions of the manufacturer (Ion PGM Sequencing 200 Kit v2) (Thermo Fisher Scientific) by Ion Personal Genome Machine (PGM) in Ion 316 Chip v2.

Phylogenetic analyses

Reads obtained by Ion Torrent sequencer were checked by quality control, cleaned up and trimmed using CLC Workbench version 5.5.1 (www.clcbio.com). A de novo assembly was performed using the default parameters and excluding contigs shorter than 1000 bases. Reads were mapped against the full genome using an online tool (Bowtie2, Galaxy Aries) and visualized by IGV software. MEGA7 was used to edit, align nucleotide and amino acid sequences and to calculate the pairwise identities of the genomes and all ORFs that were predicted using the online tool ORF Finder (NCBI, http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Complete genome sequences of MERS-CoV and apha- and beta CoVs from bats, human and camels (n. 131) were obtained from the NIAID Virus Pathogen Database and Analysis Resource (ViPR) [18] through the web site at http://www.viprbrc.org/. Multiple sequence alignment was calculated using the MUSCLE algorithm. The maximum likelihood phylogenetic tree was performed using IQtree software [19] and Model finder to determine the best model according to BIC [20]. Genetic relationships between Italian and SA bats and MERS-CoV were confirmed by comparison of the sequence distances of MERS-CoV and bat-BCoV 2c (SA bat and IT bats) using SSE v1.2 [21]. The complete S protein and the S1 domain responsible for DPP4 recognition (located in a C-terminal 240 residue RBS of IT-bat CoVs) were compared with those of MERS-CoV, HKU4, HKU5 and MERS-related bat CoVs from China and Africa. To better investigate the relationship between MERS-CoV and related bat sequences, a maximum likelihood phylogenetic tree based on the S1 protein was constructed using the IQtree software including only beta CoV sequences. Similarity plots of CoV clade c genomes were generated by using SSE v1.2 using a sliding window of 600 and a step size of 100 nucleotides (nt).

Three-dimensional macromolecular structure

The three dimensional macromolecular structure (3D–MMS) of the DPP4 binding domain in the S protein was predicted using the sequence-homology method that is based on sequences and structures released by the protein data bank (PDB) and visualized by Cn3D v4.3 software [22]. The secondary structure elements are defined based on an ESPript (http://espript.ibcp.fr) algorithm [23] and are labelled in a previous report on the MERS-RBD structure [24].

Results

Two bat carcasses belonging to two different species, *Hypsugo savii* and *Pipistrellus kuhlii*, were provided in 2011 by a wildlife recovery centre located in the Modena province (North Italy). The first bat was an adult female whereas for the second one data on age and sex were not available. During necropsy, no pathological lesions indicative of infectious diseases were observed in the two animals, but dehydration and traumatic injuries such as lacerations of the wing membrane were observed.

Samples of intestine from the two bats tested positive by the pan-coronavirus one-step RT-PCR method, and then cell cultures were inoculated with them to attempt virus isolation without success.

Genome organization, phylogenetic and molecular analyses

Two complete genome sequences of Bat-CoV/Hypsugo savii/206645–40/2011 (BatCoV-Ita1) and Bat-CoV/Pipistrellus khulii/Italy/206645–63/2011 (BatCoV-Ita2) were obtained from total RNA extracted from portions of intestine. Comparison of the RdRp sequences of the two samples obtained by Sanger method by Lelli et al. [12] showed 99% nucleotide identity between them; Initial BLAST analysis revealed they were highly similar to the beta CoVs clade 2csequences. The two full genome sizes were 30,048 nt for BatCoV-Ita1 and 30,039 for BatCoV-Ita2, with a G + C content of 39% each.

The genome organization of BatCoV-Ita1 and BatCoV-Ita2 (IT bat CoVs) is identical to that of MERS CoV species encompassing the 10 open reading frames (ORFs) in the order of ORF1ab-spike-ORF3-ORF4ab-ORF5-envelope (E)-membrane (M)-nucleocapsid (N)-ORF8b and the common non-translated sequences identified in CoV genomes at the 5′ and 3′ genomic termini and between ORF5 and the E gene (Table 1).

In ORF1ab the predicted slippery sequence "UUUAAAC" has been observed fitting the consensus motif X_XXY_YYZ (where XXX normally represents any three identical nucleotides; YYY represents strictly AAA or UUU; and Z represents A, C, or U) of nidoviruses involved in synthesis of the replicase pp1ab polyprotein by ribosomal frameshift.

The size and genomic localization of the nonstructural protein (NSP 1–16) encoded by ORF1ab were predicted by sequence comparison with other beta CoV species. Table 2 shows the 15 expected cleavage sites, 11 recognized by the "main protease" 3C–like protease (3CLpro,

Table 1 Genome localization of predicted protein sequences, putative leader TRS-L and TRS-B

BatCoV-Ita1			
ORF	nt position (start-end)	No. of amino acids	Sequence[a]
ORF1ab (TRS-L)	217–21,446	7076	$_{00036}$GATTTTAACGAACTTAAA$_{00053}$
Spike	21,388–25,425	1345	$_{21330}$C.AG.....CGTT$_{21347}$
ORF3	25,438–25,749	103	$_{25419}$TCAC.A.....T.....T$_{25436}$
ORF4a	25,758–26,045	95	$_{25742}$A.AA.....CT.T$_{25759}$
ORF4b	25,963–26,724	253	
ORF5	26,731–27,414	227	$_{26,717}$.G.GG.....ATGG$_{26734}$
E	27,493–27,741	82	$_{27479}$TTGGAA.....ATGT$_{27496}$
M	27,756–28,412	218	$_{27,734}$.GG.....CTCT$_{27751}$
N	28,460–29,749	429	$_{28,430}$.....TC.TT.$_{28447}$
ORF8b	28,506–29,084	192	
BatCoV-Ita2			
ORF	nt position (start-end)	No. of amino acids	Sequence[a]
ORF1ab (TRS-L)	208–21,437	7076	$_{00026}$GATTTTAACGAACTTAAA$_{00043}$
Spike	21,379–25,416	1345	$_{21321}$C.AG.....CGTT$_{21338}$
ORF3	25,429–25,740	103	$_{25410}$TCAC.A.....T.....T$_{25427}$
ORF4a	25,749–26,036	95	$_{25733}$A.AA.....CT.T$_{25750}$
ORF4b	25,954–26,715	253	
ORF5	26,722–27,405	227	$_{26,708}$.G.GG.....ATGG$_{26725}$
E	27,484–27,732	82	$_{27470}$TTGGAA.....ATGT$_{27487}$
M	27,747–28,403	218	$_{27,725}$.GG.....CTCT$_{27742}$
N	28,451–29,740	429	$_{28,421}$.....TC.TT.$_{28438}$
ORF8b	28,497–29,075	192	

[a]Dots represent identical nucleotides compared to the TRS-L

Table 2 Prediction of the putative pp1a/pp1b cleavage sites of BatCoV-Ita1/2 based on sequence comparison with MERS-CoV strain HCoV-EMC/2012

NSP	Position of the putative cleavage sites[a]	Protein size (no. of amino acids)	Putative functional domain(s)[b]
NSP1	Met^{1}-Gly^{195}	195	
NSP2	Asn^{196}-Gly^{855}	660	
NSP3	Ala^{856}-Gly^{2738}	1883	ADRP, PL2pro
NSP4	Ala^{2739}-Gln^{3245}	507	
NSP5	Ser^{3246}-Gln^{3551}	306	3CLpro
NSP6	Ser^{3552}-Gln^{3843}	292	
NSP7	Ser^{3844}-Gln^{3926}	83	
NSP8	Ala^{3927}-Gln^{4125}	199	Primase
NSP9	Asn^{4126}-Gln^{4235}	110	
NSP10	Ala^{4236}-Gln^{4375}	140	
NSP11	Ser^{4376}-Ile^{4389}	14	Short peptide at the end of *ORF1a*
NSP12	Ser^{4376}-Gln^{5309}	934	RdRp
NSP13	Ala^{5310}-Gln^{5907}	598	HEL, NTPase
NSP14	Ser^{5908}-Gln^{6431}	524	ExoN, NMT
NSP15	Gly^{6432}-Gln^{6773}	342	NendoU
NSP16	Ala^{6774}-His^{7076}	303	OMT

[a]Superscript numbers indicate positions in polyprotein pp1a/pp1ab or position in available sequence with the supposition of a ribosomal frameshift based on the conserved slippery sequenced (UUUAAAC) of Coronaviruses. Localized at nucleotide position 13,359–13,365 for BatCoV-Ita1 and 13,350–13,356 for BatCoV-Ita2
[b]*ADRP* ADP-ribose 1-phosphatase, *PL2pro* papain-like protease 2, *3CLpro* coronavirus NSP5 protease, *Hel* helicase, *NTPase* nucleoside triphosphatase, *ExoN* exoribonuclease, *NMT* N7 methyltransferase, *NendoU* endoribonuclease, *OMT* 2′ O-methyltransferase

NSP4–10, NSP12–16), 3 by papain-like protease (PL2pro, NSP1–3) as well as the autocatalytic site (NSP11). The IT bat CoV cleavage sites, recognized by viral proteases, were identical to those of a BatCoV isolated in China (BtVs BetaCoV/SC2013) and differed from the MERS-CoV by one amino acid in the cleavage sites between NSP1/2 and NSP6/7 (Table 3). A predicted leader transcription regulatory sequence (TRS-L), and seven putative transcription regulatory sequences body TRS-B, representing signals for the discontinuous transcription of subgenomic mRNAs (sgmRNAs), have been identified. The two IT bat isolates shared the same TRS-L, the seven TRS-B (Table 1) as well as 98.8% nucleotide identity. Across the whole genome, the percentage of overall nucleotide identity among other beta CoVs was 80% to MERS CoV, 82% to SC2013, 82% to NeoCoV, 81% to PREDICT, 72.4% to bat CoVs HKU4 and 72,5% to HKU5. The genomic sequence identity between IT bat CoVs, MERS and other MERS-related Bat CoVs is reported in Fig. 1; in particular the lowest identity of the IT bat CoVs with the other beta CoV strains was evidenced in the genomic regions encoding Pl2pro (NSP3) within ORF1ab, Spike, ORF4ab and ORF5.

A phylogenetic tree of the complete genomes showed a monophyletic group placed in the Beta2c clade formed by MERS-CoV sequences originating from humans and camels and bat-related sequences. The closest bat sequences are those originating from Africa, Italy and China (Fig. 2).

Comparison of the predicted protein sequences of the IT bat CoVs and the other beta CoVs showed the highest amino acid sequence identities for E (69.5–91.5%), M (81.7–83.7) and N (71.8–83.7%) proteins and the lowest in ORF3 (28.6–53.5%) and ORF 4b (29.9–56%)(Table 4). Comparison of MERS-CoV protein sequences to EriCoV, HKU4 or HKU5 displayed the highest amino acid sequence identities in the E, M and N proteins whereas the lowest were observed in the S and ORF4b.. In particular, the S protein of MERS-CoV showed above 50% identity to the related bat S proteins: 68.8% to IT bat CoVs, 68.6% to SC2013, 67.7% to HKU4, 64.9% to HKU5 and 64% to NeoCoV. However, if we analyse only

Table 3 Comparison of the predicted pp1a/pp1b cleavage site sequences[a] of BatCoV-Ita1/2 with prototype clade c betacoronaviruses and MERS related strains

	NSP1	NSP2	NSP3	NSP4	NSP5	NSP6	NSP7	NSP8	NSP9	NSP10	NSP11	NSP12	NSP13	NSP14	NSP15
BatCoV-Ita1/2	LVGG	LKGG	IVGG	LQS	MQS	VQS	LQA	LQN	LQA	TQS	RGSI	LQA	LQS	VQG	LQA
MERS[b]	-I–	–	–	–	–	M–	–	–	–	P–	–	–	–	–	–
HKU4[c]	-I–	–	–	–	–	–	–	–	–	P–	GS-V	–	–	–	–
HKU5[d]	–	–	LS–	–	–	–	–	–	–	P–	–	–	–	I–	–
Erinaceus[e]	-C–	–	–	–	–	–	–S	–	–	LH-	–	–	–	–	–
NeoCoV[f]	-T–	–	–	–	–	I–	–	–	–	P–	–	–	–	–	–
BtVs-BetaCoV/SC2013[g]	–	–	–	–	–	–	–	–	–	P–	–	–	–	–	–

[a]Hyphens represent identical amino acids compared to the BatCoV-Ita1/2 sequences
[b]GenBank accession number JX869059
[c]GenBank accession number EF065505
[d]GenBank accession number EF065509
[e]GenBank accession number KC545386
[f]GenBank accession number KC869678
[g]GenBank accession number KJ473821

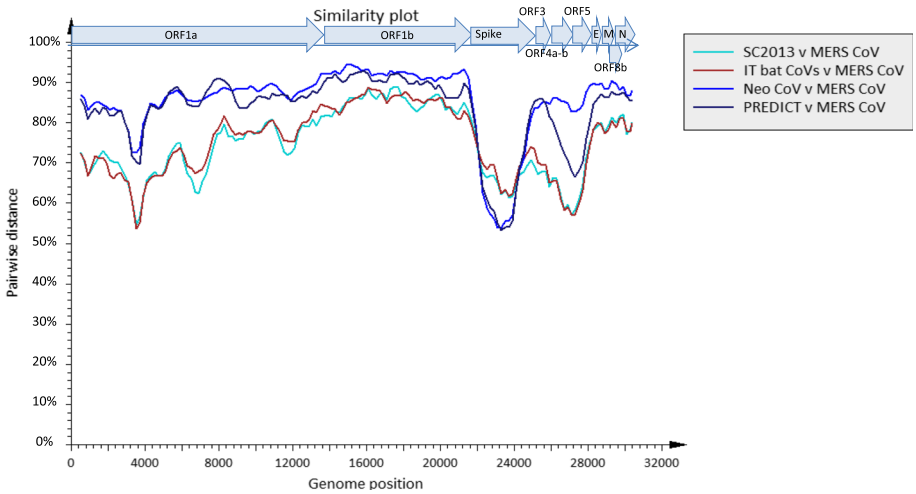


Fig. 1 Sequence identity between IT bat CoVs and other prototype clade c betacoronaviruses and MERS related strains. Similarity plots were generated using SSE version 1.2 using a sliding window of 600 and a step size of 100 nucleotides

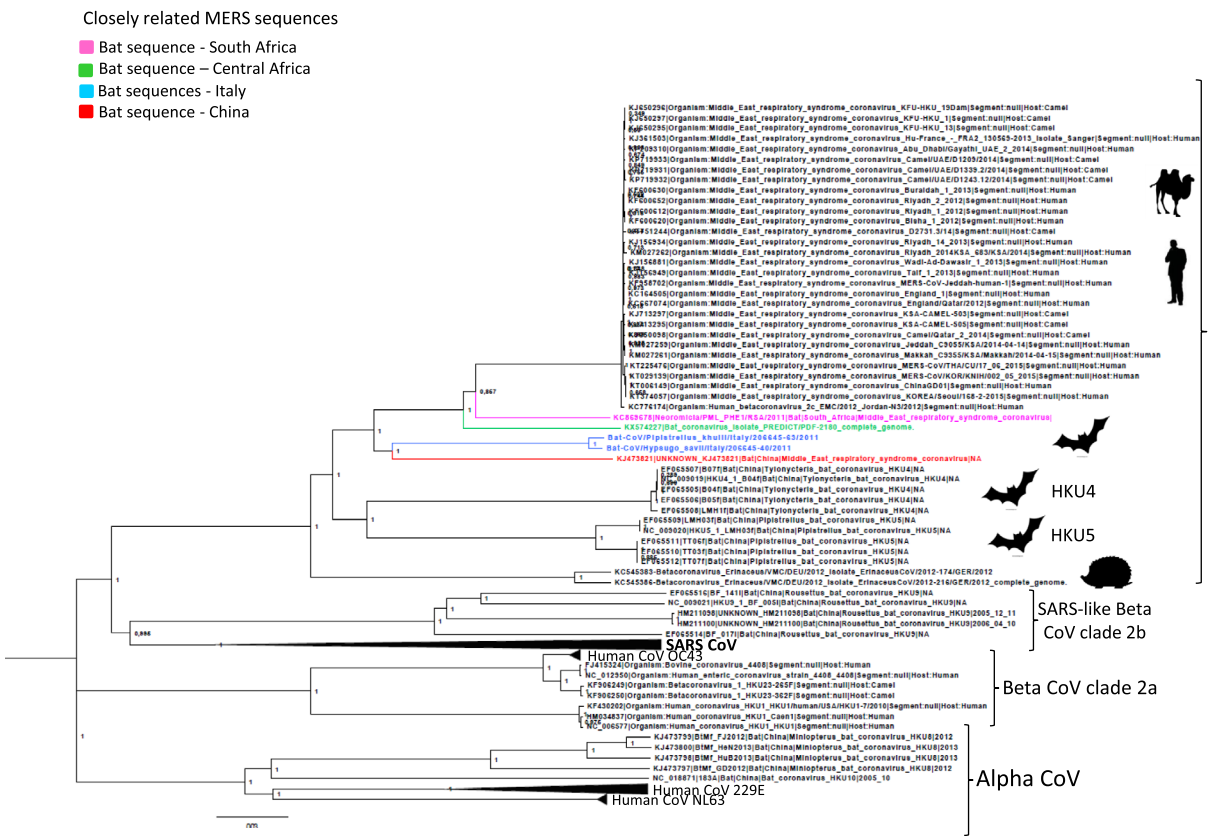


Fig. 2 Maximum phylogenetic tree based on alpha and beta CoV full genomes. Ultrafast bootstrap approximation approach was performed to compute the support of phylogenetic groups. Best-fit model according to BIC was GTR + G4. Bat sequences close related to MERS-CoVs are reported in colors: light blue for IT bat CoVs, pink for NeoCoV from South Africa, green for PREDICT from Uganda and red for SC2013 from China. Bat-CoV/H.savii/Italy/206645–40/2011 and Bat-CoV/P.khulii/Italy/206645–63/2011 sequences can be retrieved under accession numbers MG596802 and MG596803 respectively

Table 4 Comparison between predicted protein sequences of the IT bat CoVs and prototype clade c betacoronaviruses and MERS related strains

	% Amino acid identities[a]						
ORF	IT Bat CoVs	MERS-CoV[b]	HKU4[c]	HKU5[d]	EriCoV[e]	NeoCoV[f]	BtVs-BetaCoV/SC2013[g]
ORF1ab	99.4	81.3-81.5	73.5–73.6	75.9–76	74.5–74.7	81.6	84.6–84.7
Spike	99	68.5–68.8	70.4–70.6	73.8–74.3	57.8–58.1	60.9	79.6
ORF3	97.1	48.5–49.5	39.6–40.7	46.5–47.5	28.6	51.5	53.4
ORF4a	97.9	53.7	44.2	50–52.2	43.2–43.4	54.7	67
ORF4b	98	43.9–46.4	29.9–30.3	31.1–32	39.9–40.4	47.4	56
ORF5	98.7	64.3–64.7	47.6	54.5–55	52.9	62.9	74
E	100	86.6	70.7	69.5	78	91.5	91.5
M	99.5	84.9–85.8	81.7–82.1	82.1–82.6	83.9–84.4	84.9	87.2
N	99.5	81.4–81.9	73–73.4	71.8	74.4–74.6	83.7	83.4
ORF8b	96.9	63.4–67	49.2–52.4	49.5–52.1	47.9–49.5	63–65.1	64–65.6
Concatenated domains	99.3	79	74	74.6	74.4	79.2	82.3

[a]Calculated with MEGA7 using a pairwise deletion option
[b]GenBank accession number JX869059, KC164505, KC776174, KF186567, KF192507, KF600612, KF600620, KJ477102
[c]GenBank accession number EF065505, EF065506, EF065507, EF065508, DQ648794
[d]GenBank accession number EF065509, EF065510, EF065511, EF065512
[e]GenBank accession number KC545386, KC545383
[f]GenBank accession number KC869678
[g]GenBank accession number KJ473821

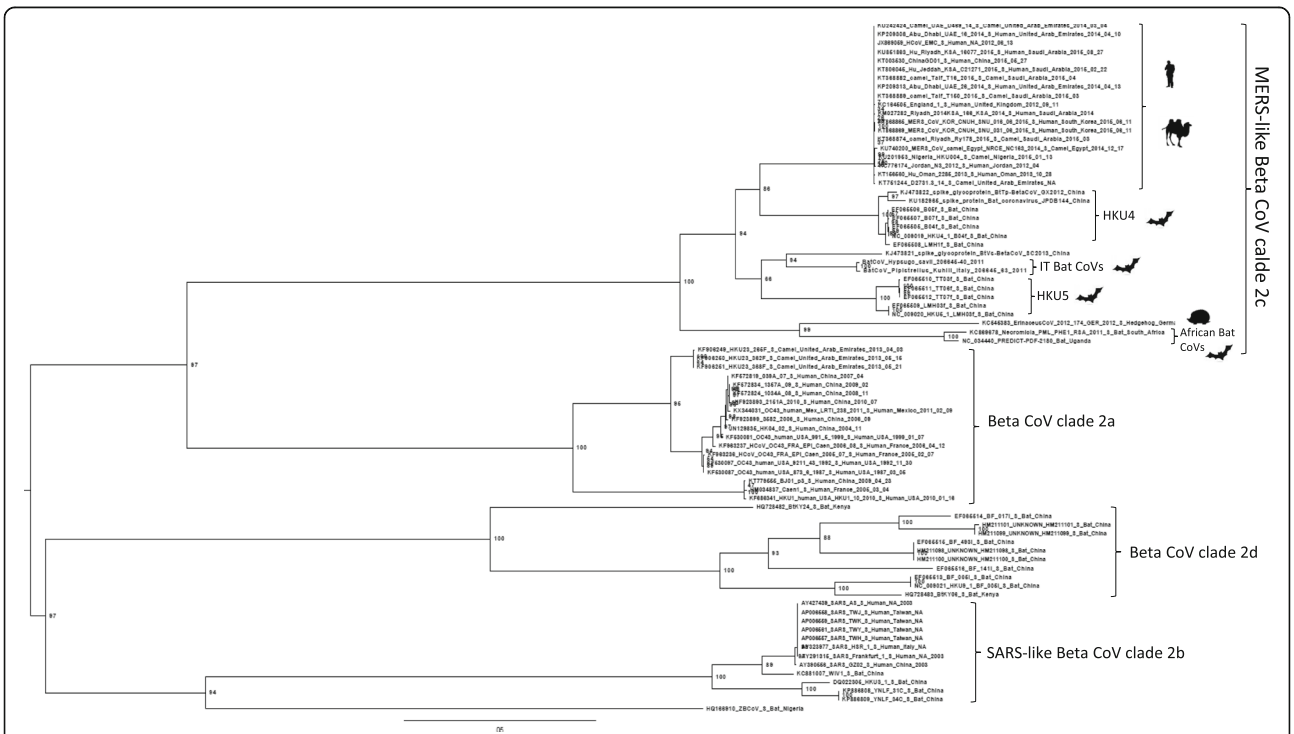


Fig. 3 Maximum phylogenetic tree based on deduced amino acid sequences of S1 protein of beta CoVs. Best-fit model according to BIC was WAG + F + I + G4. Clades 2a, 2b, 2c and 2d corresponding to the sequences of human CoVs OC43, SARS-like, MERS-like and HKU9 CoV respectively are identified

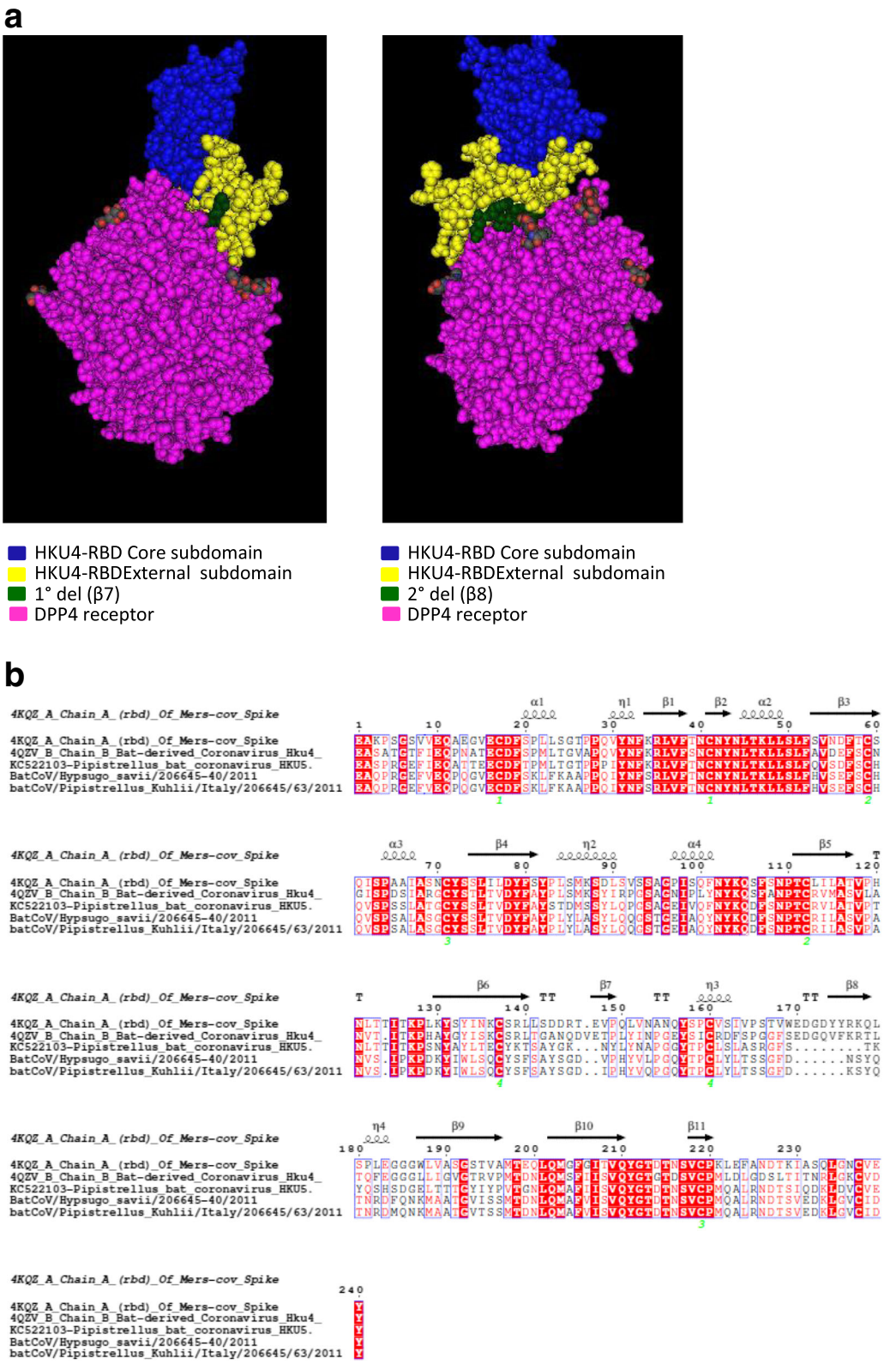


Fig. 4 (See legend on next page.)

(See figure on previous page.)
Fig. 4 Comparison of secondary and tertiary structures of MERS-like CoV RBD. **a** Predicted 3D MMS of the core and external subdomains of HKU4 RBD and human DPP4 (right and left lateral view). The two deletions observed in IT Bat CoVs RBD are evidenced in green. **b** Structure –based sequence alignment . The secondary structure elements are defined based on an ESPript algorithm and are labeled as in a previous report on the MERS RBD structure [24]. Spiral lines indicate helices, while arrows represent β strands. The external subdomain is highlighted by enclosure in a blue box. The two deletions found in IT Bat CoVs and HKU5 RBD are marked with blue lines. The Arabic numerals 1–4 indicate cysteine residues that pair to form disulfide bonds

the RBD region spanning amino acids 367–606 within the S1 subunit, which is the DPP4-interacting region, the percentage of identity to MERS changed to being slightly higher (54.1–55.0%) for HKU4 with respect to other CoVs (HKU5, IT-batCoVs) (52.5–51.7%). The RBD of Neo CoV showed the lowest percentage of identity to MERS (33.5%).

Moreover, the phylogenetic tree based on the S1 protein (Fig. 3) shows four clearly different clades: 2a, 2b, 2c and 2d corresponding to the sequences of human CoVs OC43, SARS-like, MERS-like and HKU9 CoV, respectively. Clade 2c reflects the RBD percentage of identity and was further differentiated into three groups: MERS sequences and the highly related HKU4 sequences form one group; HKU5 and IT-batCoVs sequences were placed in the second group; the third one, which is the most distant, includes African bats and hedgehog sequences.

The secondary and three dimensional structures of the RBD domain of IT bat CoVs were analysed in comparison with the MERS, HKU4 and HKU5 sequences. Two aa deletions located in a region corresponding to the external subdomain of MERS-RBD were found in IT bat CoV-RBDs in the same positions as in HKU5-RBD: three and eight aa in the HKU5-RBD and two and six in IT bat CoVs. The two deletions are located in two regions corresponding to scaffold strands β7 and β8 in the MERS/HKU4-RDB structure (Fig.4a, b). These two β strands together with β6 and 9 form the external subdomain characterized by four anti-parallel β strands that expose a flat sheet-face for receptor engagement [8].

Discussion

The high diversity of bat species as well as other unique biological and ecological features, such as long life span, roosting, migratory behaviour and the use of torpor and hibernation, contributes to bats being considered natural hosts of a large number of diverse viruses [25]. Another important characteristic is the evolution of flight, which is the most peculiar characteristic of bats and one of the most important for their wide distribution; it may have had effects on some aspects of the evolution of the immune system and the metabolism of bats and could allow them to host different viruses [26, 27]. Bats are also demonstrated natural reservoirs of many alpha CoVs and beta CoVs, which provide viral genes for the genesis of newly emerging coronaviruses with interspecies transmission potential.

Because of its similarity to the SARS CoV, it had been proposed that bats were somehow involved in transmission of the MERS CoV. Indeed, Memish et al. [5] detected a partial RNA sequence of a beta CoV obtained from a faecal pellet from an Egyptian tomb bat that showed 100% identity to the virus from the human index case-patient. The emergence of MERS CoV probably involved genetic exchanges between different viral ancestors that may have occurred either in bat ancestors or in camels acting as mixing vessels for viruses from different hosts. Recent studies have suggested that one-humped camels (*Camelus dromedarius*) may be a primary source of this virus in nature [28], and experimental infections of camels with MERS CoV seem to support this view [29].

In this study, the full genomes of two beta CoVs closely related to MERS obtained from two Italian bats belonging to the *P. kuhlii* and *H. savii* species are reported. Italy is an area of high bat species diversity with more than 30 bat species documented by historical records and recent studies, but these two species, which belong to Vespertionilidae family, are the most frequently recorded [30]. *Pipistrellus kuhlii* forages over a variety of habitats, including agricultural and urban areas (including around street lights). Recent evidence suggests that urbanization may be beneficial to this species in that colonies in urban and suburban areas have advanced parturition and produce more offspring than colonies in rural areas, at least in central Italy [31]. *H. savii* forages over open woodland, pasture and wetlands and often feeds at lights in rural areas, towns and cities. This is one of the most common species in the Italian Mountains, the Apennines and the Alps below 2600 m.

Detection of viruses belonging to clade 2c seems to be particularly associated with vespertilionid bats even if this association is not exclusive. Indeed, NeoCoV, BtVs-BetaCoV/SC2013, PREDICT/PDF-2180, HKU4, HKU5 and the two IT bat CoVs were all found in species belonging this family.

The full IT bat CoV sequences were obtained from two bat carcasses and showed the same genome organization as MERS-CoV either for the 10 open reading frames (ORFs) or the common non-translated sequences identified in CoV genomes. The overall nucleotide identity

to MERS CoV is close to 78%, and in the phylogenetic tree they are represented in the same MERS-like clade 2c. From the molecular point of view, the International Committee on Taxonomy of Viruses (ICTV) has established 90% amino acid sequence identity as the threshold value for CoV species demarcation of the seven concatenated domains within the ORF1ab: NSP3 (ADRP), NSP5 (3CLpro), NSP12 (RdRp), NSP13 (Hel, NTPase), NSP14 (ExoN, NMT), NSP15 (NendoU), NSP16 (OMT). The sequence identity of the BatCoV-Ita1 and BatCoV-Ita2 concatenated domains is below the threshold value compared to HKU4, HKU5 and EriCoV (86.1–89.2%) and over the threshold value compared to MERS-CoV, NeoCoV and BtVs-BetaCoV/SC2013 (92–92.9%), indicating that the two IT bat CoVs could be included in the same virus species as MERS-CoV and related isolates.

Full genome phylogenetic reconstruction showed that the two African bat-CoV sequences were the ones most closely related to MERS; however, the spike gene evidenced higher sequence differences with respect to HKU4 and other related bat sequences, IT bat CoVs included. Indeed, it was demonstrated that the RBD domain in the S1 protein of HKU4 but not HKU5 can bind to the human DPP4 receptor even if with less affinity. The marked difference between HKU4-RBD and HKU5-RBD with respect to MERS-RBD is the presence of two marked deletions in the external subdomain responsible for receptor recognition [24]. Two similar deletions in the region corresponding to scaffold strands β7 and β8 in the MERS/HKU4-RBD structure were observed in IT bat CoVs, suggesting also for them the absence of DPP4-binding potential. Based on these results we can hypothesize that human DPP4 is not a functional receptor for IT bat CoVs as previously shown for HKU5-CoVs.

Conclusions

The role played by bats in the maintenance and transmission of beta CoVs, if they are simply incidental hosts or competent reservoir hosts able to transmit them to other vertebrates, is an open question that must be carefully addressed. It is believed that the majority of all alpha and beta CoVs currently circulating in mammals are evolutionarily linked to ancestral CoVs originated from bats [4]. However, more surveillance studies are needed to better investigate the potential intermediate hosts that may play a role in the interspecies transmission of known and currently unknown coronaviruses. Particular attention should be paid to investigating the S protein sequences and structures as well as receptor specificity and binding affinity as keys to understanding the biology of bat-derived viruses, their potential threat to human health and the evolutionary pathway of MERS-CoV.

Abbreviations

3DMMS: Three dimensional macromolecular structure; CoV: Coronavirus; DPP4: Dipeptidyl peptidase 4; ICTV: International Committee on Taxonomy of Viruses; MERS-CoV: Middle Est respiratory syndrome coronavirus; PDB: Protein data bank; RBD: Receptor binding domain; SARS-CoV: Severe acute respiratory syndrome coronavirus; TRS-B: Transcription regulatory sequence body; TRS-L: Transcription regulatory sequence leader

Acknowledgements

We thank Enrica Rosti and Piero Milani of the Centro Fauna Selvatica "Il Pettirosso" of Modena (Italy) for their precious assistance with sampling and Francesca Faccin and Marta Consoli for their technical collaboration.

Funding

This study was partially funded by the Italian Ministry of Health under the Research projects: GR-2011-02350591 PGR2011001- "An epizootiological survey of bats as reservoirs of emerging zoonotic viruses in Italy: implications for public health and biological conservation"; ISS PE-2011-02351681 PRF2011301 "Emerging respiratory viruses: monitoring of coronavirus infections at the human-animal interface". The funding body was not involved into the design of the study, and collection, analysis, and interpretation of data in the manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

AM, DL and GV participated in the study conception and design. AM, GV drafted the manuscript DL, AP, ES: conducted sampling of bat fecal samples and bat species identification LS, GZ, AB, ES, AP, EC, AL, AM: conducted full genome sequencing and participated in molecular genetic studies MRC, MC: performed critical revision All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Department of Virology, Istituto Zooprofilattico Sperimentale Lombardia ed Emilia Romagna, Via Antonio Bianchi 9, 25124 Brescia, Italy. [2]Department of Food Safety, Istituto Superiore di Sanità. Nutrition and Veterinary Public Health, Viale Regina Elena 299, 00161 Rome, Italy. [3]Dept. of Sciences, University Roma Tre, Viale Guglielmo Marconi 446, 00146 Rome, Italy. [4]University Campus Bio-Medico of Rome, Via Álvaro del Portillo, 21, 00128 Rome, Italy.

Received: 20 October 2017 Accepted: 6 December 2017
Published online: 19 December 2017

References

1. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, genetic recombination and pathogenesis of coronaviruses. Trends in Microbiol. 2016;26:490–502. https://doi.org/10.1016/j.tim.2016.03.003.
2. WHO. http://www.who.int/emergencies/mers-cov/en/.
3. Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nature Rev Microbiol. 2016;14:523–34. https://doi.org/10.1038/nrmicro.2016.81.
4. Hu B, Ge X, Wang LF, Shi Z. Bat origin of human coronaviruses. Virol J. 2015; 12:221. https://doi.org/10.1186/s12985-015-0422-1.
5. Memish ZA, Mishra N, Olival KJ, Fagbo SF, Kapoor V, Epstein JH, et al. Middle East respiratory syndrome coronavirus in bats. Saudi Arabia Emerg Infect Dis. 2013;19(11):1819–23. https://doi.org/10.3201/eid1911.131172.

6. Anthony SJ, Gilardi KV, Menachery D, Goldstein T, Ssebide B, Mbabazi R, et al. Further evidence for bats as the evolutionary source of middle east respiratory syndrome coronavirus. MBio. 2017;8(2):e00373–17. https://doi.org/10.1128/mBio.00373-17.
7. Raj VS, Mou H, Smits SL, Dekkers DH, Müller MA, Dijkman R, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature. 2013;495(7440):251–4. https://doi.org/10.1038/nature12005.
8. Wang Q, Qi J, Yuan Y, Xuan Y, Han P, Wan Y, et al. Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. Cell Host Microbe. 2014;16(3):328–37. https://doi.org/10.1016/j.chom.2014.08.009.
9. Munster VJ, Adney DR, van Doremalen N, Brown VR, Miazgowicz KL, Milne-Price S, et al. Replication and shedding of MERS-CoV in Jamaican fruit bats (Artibeus Jamaicensis). Sci Rep. 2016;22(6):21878. https://doi.org/10.1038/srep21878.
10. Annan A, Baldwin HJ, Corman VM, Klose SM, Owusu M, Nkrumah EE, et al. Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. Emerg Infect Dis. 2013;19(3):456–9. https://doi.org/10.3201/eid1903.121503.
11. Anthony SJ, Ojeda-Flores R, Rico-Chávez O, Navarrete-Macias I, Zambrana-Torrelio CM, Rostal MK, et al. Coronaviruses in bats from Mexico. J Gen Virol. 2013;94(5):1028–38.
12. Lelli D, Papetti A, Sabelli C, Rosti E, Moreno A, Boniotti MB. Detection of coronaviruses in bats of various species in Italy. Viruses. 2013;5(11):2679–89. https://doi.org/10.3390/v5112679.
13. Ithete NL, Stoffberg S, Corman VM, Cottontail VM, Richards LR, Schoeman MC, et al. Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. Emerg Infect Dis. 2013;19(10):1697–9. https://doi.org/10.3201/eid1910.130946.
14. Corman VM, Ithete NL, Richards LR, Schoeman MC, Preiser W, Drosten C, et al. Rooting the phylogenetic tree of Middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. J Virol. 2014;88(19):11297–303.
15. Yang L, Wu Z, Ren X, Yang F, Zhang J, He G, et al. MERS–related Betacoronavirus in Vespertilio Superans bats. China Emerg Infect Dis. 2014;20(7):1260–2. https://doi.org/10.3201/eid2007.140318.
16. Dietz C, von Helversen O. Illustrated identification key to the bats of Europe. Version 1.0, First released 15.12.2004. https://rhone-alpes.lpo.fr/images/chiroptere/telecharger/dietz_von_helversen_2004_1.pdf.
17. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, et al. Viral genome sequencing by random priming methods. BMC Genomics. 2008;9:5. https://doi.org/10.1186/1471-2164-9-5.
18. Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, et al. ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res. 2012;40(Database issue):D593–8. https://doi.org/10.1093/nar/gkr859.
19. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol Biol Evol. 2015;32:268–74. https://doi.org/10.1093/molbev/msu300.
20. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods. 2017;14:587–9. https://doi.org/10.1038/nmeth.4285.
21. Simmonds PSSE. A nucleotide and amino acid sequence analysis platform. BMC Res Notes. 2012;20(5):50. https://doi.org/10.1186/1756-0500-5-50.
22. Wang Y, Geer LY, Chappey C, Kans JA, Bryant SH. Cn3D: sequence and structure views for Entrez. Trends Biochem Sci. 2000;25:300–2. https://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml.
23. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucl Acids Res. 2014;42(W1):W320–4. https://doi.org/10.1093/nar/gku316.
24. Lu G, Hu Y, Wang Q, Qi J, Gao F, Li Y, et al. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature. 2013;500(7461):227–31. https://doi.org/10.1038/nature12328.
25. Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. Bats: important reservoir hosts of emerging viruses. Clin Microbiol Rev. 2006;19(3):531–45. https://doi.org/10.1128/CMR.00017-06.
26. O'Shea TJ, Cryan PM, Cunningham AA, Fooks AR, Hayman DTS, Luis AD, et al. Bat flight and zoonotic viruses. Emerg Infect Dis. 2014;20(5):741–5. https://doi.org/10.3201/eid2005.130539.
27. Brook CE, Dobson AP. Bats as 'special' reservoirs for emerging zoonotic pathogens. Trends Microbiol. 2015;23(3):172–80. https://doi.org/10.1016/j.tim.2014.12.004.
28. Raj VS, Farag EABA, Reusken CBEM, Lamers MM, Pas SD, Voermans J, et al. Isolation of MERS coronavirus from a dromedary camel, Qatar, 2014. Emerg Infect Dis. 2014;20(8):1339–42. https://doi.org/10.3201/eid2008.140663.
29. Adney DR, van Doremalen N, Brown VR, Bushmaker T, Scott D, de Wit E, et al. Replication and shedding of MERS-CoV in upper respiratory tract of inoculated dromedary camels. Emerg Infect Dis. 2014;20(12):1999–2005. https://doi.org/10.3201/eid2012.141280.
30. Russo D, Jones G. Use of foraging habitats by bats in a Mediterranean area determined by acoustic surveys: conservation implications. Ecography. 2003;26:197–209. https://doi.org/10.1034/j.1600-0587.2003.03422.
31. Ancillotto L, Santini L, Ranc N, Maiorano L, Russo D. Extraordinary range expansion in a common bat: the potential roles of climate change and urbanization. Naturwissenschaften. 2016;103(3–4):15. https://doi.org/10.1007/s00114-016-1334-7.

CHAPTER 8

Full genome characterization of two novel Alpha-coronavirus species from Italian bats

8. Full genome characterization of two novel Alpha-coronaviruses species from Italian bats

Since the SARS and MERS outbreaks, the surveillance and phylogenetic analysis on CoV strain from bats have increased. Among Coronaviruses, not only SARS-CoV and MERS-CoV have origin from bats. Phylogenetic studies have recently showed that the two Alpha-CoVs, HCoV-229E and HCoV-NL63, have relatives with bats strains. The implementing of surveillance on bats has led to the disCoVery of novel bat CoV species related to both human and animal strains. Bat strains related to Alpha-CoV human strains have been detected in Kenyan and American bats. The phylogenetic analysis has showed that human and bat strains share a common ancestor and that bats may be the potential reservoir host also for Alpha-CoV human strains. In addition, novel and unknown CoVs species may exist since a minor percentage (20%) of the bat species have been analysed for the presence of CoVs. Despite this, analysis has confirmed the high heterogeneity of CoVs in bats. In the present study, three CoV strains were fully sequenced by NGS. The sequence analysis showed the detection of three Alpha-CoVs detected from Italian *Pipistrellus kuhlii* bats. Phylogenetic analysis on these strains showed low genetic correlation with the other Alpha-CoV species. Based on the ICTV classification criteria, these strains were classified into two novel CoV species never detected before. In addition, the Italian strains were highly divergent from human CoV strains. The phylogenetic analysis helped with the classification into novel CoV species. However, the increasing of surveillance in the future will probably lead to the discovery of novel and unknown species.

Virus Research 260 (2019) 60–66

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Full genome characterization of two novel Alpha-coronavirus species from Italian bats

Luca De Sabato[a,b], Davide Lelli[c], Francesca Faccin[c], Sabrina Canziani[c], Ilaria Di Bartolo[b], Gabriele Vaccari[b,*], Ana Moreno[c]

[a] Department of Sciences, University Roma Tre, Viale Guglielmo Marconi 446, 00146, Rome, Italy
[b] Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161, Rome, Italy
[c] Department of Virology, Istituto Zooprofilattico Sperimentale Lombardia ed Emilia Romagna, Via Antonio Bianchi 9, 25124, Brescia, Italy

ARTICLE INFO

Keywords:
Bats
Full genome sequencing
Italy
Alpha-CoV viruses

ABSTRACT

Coronaviruses (CoVs) have been detected worldwide in several bat species, which are considered the main reservoir. The attention to the high diversity of CoVs hosted by bats has increased during the last decade due to the high number of human infections caused by two zoonotic Beta-CoVs, SARS-CoV and MERS-CoV, that cause several respiratory diseases. Among coronaviruses, two Alpha-CoV strains (HuCoV-229E and HuCoV-NL63) cause mild respiratory disease that can change to severe disease in children, elderly and individuals affected by illnesses. Phylogenetic analysis conducted on bat Alpha-CoV strains revealed their evolutive correlation to human strains, suggesting their origin in bats. The genome of CoVs is characterized by a high frequency of mutations and recombination events, increasing their ability to switch hosts and their zoonotic potential. In this study, three strains of Alpha-CoV genera detected in Italian bats (*Pipistrellus kuhlii*) were fully sequenced by Next Generation Sequencing (NGS) and characterized. The complete genome analysis showed the correlation of the Italians strains with a Chinese strain detected in 2013 and, based on CoV molecular species demarcation, two new Alpha-CoV species were established. The analysis of a fragment of the *RNA-dependent RNA polymerase* (*RdRp*) showed the correlation of the Italian strains with CoVs that was only detected in the bat *Pipistrellus* genera (*Pipistrellus kuhlii* and *Pipistrellus Pipistrellus*) in European countries.

1. Introduction

Bats are considered the natural reservoirs of several emerging and re-emerging viruses, such as Nipah virus, Marburg virus, rabies virus and coronaviruses, that have caused outbreaks in both humans and animals (Shi, 2013; Smith and Wang, 2013). The ecological features of bats, including their ability to fly long distances, their longevity, their large social colonies and their potential interactions with humans or livestock animals, facilitate virus maintenance and transmission, increasing the risk of intraspecies or interspecies jumping (Calisher et al., 2006). Among bat viruses, in the last decade, a large diversity of coronaviruses has been detected, exceeding the diversity seen in other mammalian hosts (Drexler et al., 2014). Coronaviruses (CoVs) (order *Nidovirales*, family *Coronaviridae*, subfamily *Coronavirinae*) are enveloped viruses characterized by a positive-sense single-stranded RNA genome of approximately 26–32 kilobases and classified into four genera (Weiss and Leibowitz, 2011). Alphacoronavirus (Alpha-CoV) and Betacoronavirus (Beta-CoV) infect several mammal species, including humans, bats and pigs, while Gammacoronavirus (Gamma-CoV) and Deltacoronavirus (Delta-CoV) infect birds, wild felines, pigs and some marine mammal species (Woo et al., 2009b, 2012). The genome of coronaviruses is characterized by high frequency recombination and a high mutation rate, which increases their potential for interspecies and intraspecies jumping (Lai, 1992; Holmes, 2009). Six CoV strains are recognized to infect humans. Two Alphacoronaviruses (HuCov-229E, -NL63) and two Betacoronaviruses (HuCoV−OC43, -HKU1) are responsible for the common cold and severe respiratory pathologies in infants, elderly people and immunocompromised patients and are characterized by human-to-human transmission (Hu et al., 2015). The other two Betacoronaviruses species, the Severe Acute Respiratory Syndrome virus (SARS-CoV in 2002–2003) and the Middle East Respiratory Syndrome virus (MERS-CoV in 2012) caused severe respiratory pathologies with case fatality rates of 9% and 35%, respectively (WHO, www.who.int). Phylogenetic analysis on strains

* Corresponding author at: Istituto Superiore di Sanità, Department of Food Safety, Nutrition and Veterinary Public Health, Viale Regina Elena 299, 00161, Rome Italy.
E-mail address: gabriele.vaccari@iss.it (G. Vaccari).

https://doi.org/10.1016/j.virusres.2018.11.007
Received 25 September 2018; Received in revised form 7 November 2018; Accepted 13 November 2018
Available online 15 November 2018
0168-1702/

detected in bats, humans and other mammals suggested that the origin of these CoVs was in bats. The Rhinolophus bat species are considered the main reservoir for SARS-related CoVs. Bat MERS-related CoVs were also detected in African, Chinese and Italian bats (Annan et al., 2013; Ithete et al., 2013; Lau et al., 2013; Corman et al., 2014; Moreno et al., 2017), supporting the hypothesis of the bat origin. In three recent studies, related strains of the HuCoV-229E were detected in Hipposideros bats and strains of HuCoV-NL63 were detected in the American tricoloured bat (*Perimyotis subflavus*) and Kenyan *Triaenops afer* species, suggesting bats as potential reservoir host of Alphacoronavirus human strains (Pfefferle et al., 2009; Huynh et al., 2012; Corman et al., 2015). In addition, relatives of HuCoV-NL63 can be grown in immortalized bat cell lines, suggesting their potential association with bats (Huynh et al., 2012). This has led to speculations about an evolutionary origin of all mammalian CoVs in bat hosts (Woo et al., 2009a, c). However, how humans become exposed to remote wildlife viruses is not always clear (Wolfe et al., 2007).

In Europe, several studies described the presence of CoVs in bat populations detecting both Alpha-CoVs and Beta-CoVs in Germany, Spain, Luxembourg, Italy, The Netherlands, the United Kingdom, France and Hungary (Gloza-Rausch et al., 2008; Reusken et al., 2010; Falcon et al., 2011; August et al., 2012; Lelli et al., 2013; Kemenesi et al., 2014; Goffard et al., 2015; Monchatre-Leroy et al., 2017; Pauly et al., 2017) from more than 20 different bat species. The detection of the same CoV strains (100% nucleotide identity) in different colonies of the same bat species or the circulation of different genera of CoVs (Alpha-CoVs and Beta-CoVs) in the same bat species confirm the high heterogeneity of CoVs in bats and that bat-CoV diversity depends more on the species-specificity than the geography and sampling location.

However, these studies were based on the analysis of a fragment of the *RNA-dependent RNA polymerase* (*RdRp*) gene that allows the assignment of the strains to the genera and not to the species. The International Committee on Taxonomy of Viruses (ICTV) established a molecular demarcation method for species assignment using the conserved domains of replicase polyprotein and the pairwise amino acid distance of 90% as threshold value. The Alpha-CoVs are classified into 11 species, 6 of which detected in bats: *Miniopterus bat coronavirus 1*, *Bat coronavirus CDPHE15*, *Miniopterus bat coronavirus HKU8*, *Rhinolophus bat coronavirus HKU2*, *Bat coronavirus HKU10*, and *Scotophilus bat coronavirus 512*, and some strains that to date are not assigned. However, the number of bat species that host CoVs is still unknown and increases proportionally with the increasing of surveillance. In this study, we describe the full genome sequencing by Next Generation Sequencing (NGS), the characterization and the classification of two novel Alpha-CoV species detected from three Italian *Pipistrellus kuhlii* bats (Lelli et al., 2013).

2. Materials and methods

2.1. Sampling

Two bat faecal samples and one carcass from three bat *Pipistrellus kuhlii* species were provided by a rehabilitation centre from Northern Italy between 2010 and 2015 and the bats species were identified according to the European bat identification keys based on their morphologic characteristics. Faecal and organ samples positive for Alpha-CoV genera by a pan-coronavirus one-step RT-PCR (Lelli et al., 2013) were chosen for NGS analysis.

2.2. Whole-genome sequencing

Libraries were prepared following the sequence independent single primer amplification method (SISPA) (Djikeng et al., 2008). The RNA, extracted as previously described by Lelli et al. (2013), was retrotranscribed using the SuperScript IV Reverse Transcriptase (Invitrogen, Monza, Italy), starting with 9 µl of RNA and following the manufacturer's instructions. Twenty microlitres of cDNA were used to synthesize the second strand of cDNA by DNA Polymerase I Large (Klenow) Fragment (Promega, Milan, Italy) and then amplified by the Expand High Fidelity PCR System (Sigma Aldrich S.R.L., Milan, Italy). The PCR amplicons were purified using one volume of Agencourt AMPure XP beads (Beckman, Milan, Italy) following the manufacturer's instructions and eluted in 40 µl of nuclease-free water. Five hundred nanograms of purified DNA, quantified with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Rodano, MI, Italy) were digested with the EcoRV enzyme (New England BioLabs, Pero, MI, Italy) and then purified with a 1.8x volume of Agencourt AMPure XP beads (Beckman, Milan, Italy). The libraries were prepared by NEBNext Fast DNA Library Prep Set for Ion Torrent following the standard protocol for 100 ng of DNA. The barcoded libraries were mixed, and the pool was used for the Emulsion PCR performed by the Ion PGM Hi-Q OT2 Kit. The sequencing run was performed according to the manufacturer's instructions (Ion PGM Hi-Q Sequencing Kit) (Thermo Fisher Scientific) by Ion Personal Genome Machine (PGM) on the Ion 318 Chip v2.

2.3. Genome structure and phylogenetic analyses

NGS, previously described by Moreno et al. (2017) was applied in order to obtain the complete genome. Data obtained by the Ion Torrent sequencer were analysed by the online portal Galaxy Aries (https://aries.iss.it). The reads were checked, cleaned up, and trimmed, and the sequences shorter than 50 nt were filtered. Host sequences were removed by mapping the reads against the Megabat and Microbat complete genomes downloaded from Genome Browser (https://www.genome.ucsc.edu) using the Bowtie 2 tool. The reads aligned by the BLASTn tool to the bacterial non-redundant nucleotide database RefSeq (https://www.ncbi.nlm.nih.gov/refseq/; E-value $> 10^{-05}$) were removed by Galaxy Aries. The remaining reads were aligned with the viral non-redundant nucleotide database RefSeq (https://www.ncbi.nlm.nih.gov/refseq) and were parsed with the MEGAN6 software. The reads that showed no significant hits to the reference database were assigned to the unclassified reads.

The sequences classified into the *Coronaviridae* family were extracted and assembled into contigs by a *de novo* assembling method, using the default parameters, and excluding those shorter than 1000 bases using SPAdes tool (Galaxy Aries). The closest viral sequences were chosen by a BLASTn analysis and used to map the reads by the online tool Bowtie2 (Galaxy Aries). The output was visualized by the Integrative Genomics Viewer (IGV) software (http://software.broadinstitute.org/software/igv/), and the consensus sequence was extracted. Nucleotide and amino acids sequences were aligned, and the pairwise identity values were calculated with MEGA7 software (www.megasoftware.net). The open reading frames (ORFs) were predicted using the online tool ORF Finder (NCBI, http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The potential cleavage sites in the orf1ab polyprotein were predicted by amino acid sequence alignment with other CoV strains and by using the online tool NetCorona 1.0 Server (http://www.cbs.dtu.dk/services/NetCorona/) (Kiemer et al., 2004). Comparison of the sequence distances of BatCoV-Ita4 and the closest Alpha-CoV sequences were confirmed using SSE v1.2 (Simmonds, 2012).

A dataset of complete genome references for Alpha-CoVs and Beta-CoVs species from the ICTV taxonomy report (https://talk.ictvonline.org) were obtained including the full genome sequences that displayed the highest nucleotide similarity to strains sequenced in this study, which resulted in 59 CoV sequences. A second dataset was used to build a ML tree using the partial sequence of the RdRp gene (409 nt) sequenced worldwide, excluding identical strains from the same study and bat species and resulting in 226 CoVs sequences. In both ML trees Beta-CoVs from different species were used as an outgroup.

To test the presence of recombination by RDP4 (Martin et al., 2015), six different methods were applied: GENECONV, BootScan, MaxChi, Chimaera, 3Seq, and SiScan, using the default settings.

The Maximum likelihood (ML) phylogenetic trees were built using MEGA7 software, applying, as a substitution model, a general time-reversible (GTR) model with a gamma-distributed (G) rate variation across sites, a proportion of invariant sites (I) (GTR + G + I) and a bootstrap analyses of 1000 pseudo-replicates.

The Bayesian phylogenetic trees were carried out using MrBayes ver. 3.1.2 (Huelsenbeck and Ronquist, 2001) using the sequences of the predicted proteins and excluding the most divergent strains. The Metropolis-coupled Markov chain Monte Carlo (MCMC) was used, starting from a random tree, run for 500 thousand heuristic search generations, sampling every 1000 generations and discarding 25% of the samples as burn-in.

Analysis of the protein families of spike proteins and the prediction of the secondary structure were performed by the online tools: PFAM, InterProScan, the TMHMM program (http://www.cbs.dtu.dk/services/TMHMM/) (Apweiler et al., 2001; Bateman et al., 2002), Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) and the Swiss model (https://swissmodel.expasy.org).

3. Results

3.1. NGS data analysis

The NGS run produced approximately 4 million reads. A total of 1,376,444 reads were obtained for Bat-CoV/P.kuhlii/Italy/206645-41/2011 (BatCoV-Ita3). Of these reads, 8% were host reads, 78% were bacterial reads and 84,069 (6%) were viral reads, of which 81,069 (5.8%) were classified as *Coronaviridae* and 8% were unclassified. For Bat-CoV/P.kuhlii/Italy/3398-19/2015 (BatCoV-Ita4), 970,190 reads were retrieved, of which 6.8% were host reads, 38% were bacterial, 72,127 (7.4%) were viral, of which 5484 (0.5%) were *Coronaviridae,* and 52% were unclassified. For Bat-CoV/P.kuhlii/Italy/206679-3/2010 (BatCoV-Ita5), 1,602,274 reads were obtained. Of these, 0.3% were host reads, 98% from bacteria, and 22,621 were viral (1,4%), of which 19,053 (1.1%) were *Coronaviridae* sequences and 0.3% were unclassified.

The reads classified into the *Coronaviridae* family were used to assemble the contigs obtaining 1 contig of approximately 27,000 nt for BatCoV-Ita3, 6 contigs > 4000 nt for BatCoV-Ita4 and BatCoV-Ita5 contigs > 2800 nt for BatCoV-Ita5. The three assembled full genomes showed an average coverage of 751x, 50x and 181x, for BatCoV-Ita3, BatCoV-Ita4, and BatCoV-Ita5, respectively. The RdRp sequence of BatCoV-Ita4 was not obtained by the Sanger method used by Lelli et al. (2013). The other two Italian bat RdRp sequences showed 99% nucleotide identity with the RdRp region of the complete genomes obtained by NGS.

3.2. Genome organization

The complete genome sizes were 27,862 nt for BatCoV-Ita3, 28,129 nt for BatCoV-Ita4 and 28,146 BatCoV-Ita5, with a G + C content of 42%, 40.3% and 40.4%, respectively. The first sequence analysis was performed by BLASTn, comparing the Italian strains with those available online (https://www.ncbi.nlm.nih.gov). The BLASTn search showed similarities with an unclassified strain BtNv-AlphaCoV/SC2013 (KJ473809), and with those viruses classified into HKU10 bat and Porcine epidemic diarrhoea virus (PEDV) species. BatCoV-Ita4 and BatCoV-Ita5 shared a 97% nucleotide identity (nt. id.), and BatCoV-Ita3 shared 70.9% and 71% nt. id. with BatCoV-Ita4 and BatCoV-Ita5, respectively, at the full genome level. The same differences were observed when the ORF nucleotide sequences were aligned separately. BatCoV-Ita4 and BatCoV-Ita5 shared > 97% nt. id. in all the ORFs. BatCoV-Ita3 showed the highest differences in the S, ORF3 and N genes with < 65% nt. id. Fewer differences were observed at the ORF1ab, M and E genes (> 70% nt. id.) compared to the other two Italian strains. Their genome organization was similar to other Alpha-CoV species, comprehending 6 ORFs and two non-translated termini in the order of 5′ terminus-ORF1ab-spike-ORF3-envelope (*E*)-membrane (M)-nucleocapsid (*N*)-3′ terminus (Table 1). In the ORF1ab, it has been observed that the predicted slippery sequence "UUUAAAC" is involved in the synthesis of the replicase pp1ab polyprotein by ribosomal frameshift, a characteristic of the *Nidovirales* order. The sizes, the genomic localization and the 15 expected cleavage sites of the nonstructural protein (NSP 1–16) that are encoded by ORF1ab, were predicted by sequence comparison with other Alpha-CoV species (Table 2). BatCoV-Ita4 and BatCoV-Ita5 showed the same sequences of cleavage sites. BatCoV-Ita3, compared to the other 2 strains, showed two amino acid changes in the cleavage sites between NSP1/NSP2 and NSP12/NSP13. A leader predicted

Table 1
Locations of predicted ORFs, protein sequences, putative leader TRS-L and TRS-B.

ORF	nt position (start-end)	No. of amino acids	Sequence[a]
BatCoV-Ita3			
ORF1ab (TRS-L)	296-20166	6623	00067CTAAAC00073
Spike	20153-24274	1373	20100−−−−T−20106
ORF3	24274-24924	216	24230−−−−−−24236
E	24947-25174	75	24939−−−−−−24945
M	25180-25863	227	25170−−−−−−25176
N	25870-27174	434	25859−−−−−T25865
BatCoV-Ita4			
ORF1ab (TRS-L)	281-20175	6627	00051CTAAAC00057
Spike	20172-24371	1399	20161−−C−−−20167
ORF3	24371-25024	217	
E	25103-25330	75	25077−−−−T−25083
M	25,343-26050	235	25326−−−−−−25332
N	26058-27356	432	26046−−−−−−26052
BatCoV-Ita5			
ORF1ab (TRS-L)	253-20192	6645	00023CTAAAC00029
Spike	20189-24388	1399	20178−−C−−−20184
ORF3	24388-25041	217	
E	25120-25347	75	25094−−−−T−25100
M	25360-26067	235	25,343−−−−−−25349
N	26075-27373	432	26063−−−−−−26069

[a] Dashes represent identical nucleotides compared to the leader TRS.

Table 2
Prediction of the putative pp1ab cleavage sites.

NSP	BatCoV-ITA3	BatCoV-ITA4	BatCoV-ITA5	Putative functional domain (s)[a]
NSP1	M^{1}-A^{107}	M^{1}-G^{107}	M^{1}-G^{107}	
NSP2	G^{108}-G^{771}	G^{108}-G^{771}	G^{108}-G^{771}	
NSP3	G^{772}-G^{2361}	G^{772}-G^{2374}	G^{772}-G^{2389}	ADRP, PL2pro
NSP4	G^{2362}-Q^{2838}	G^{2375}-Q^{2851}	G^{2390}-Q^{2866}	
NSP5	A^{2839}-Q^{3140}	A^{2852}-Q^{3153}	A^{2867}-Q^{3168}	3CLpro
NSP6	S^{3141}-Q^{3419}	S^{3154}-Q^{3432}	S^{3169}-Q^{3447}	
NSP7	S^{3420}-Q^{3502}	S^{3433}-Q^{3515}	S^{3448}-Q^{3530}	
NSP8	$S3^{503}$-Q^{3697}	S^{3516}−Q^{3710}	S^{3531}-Q^{3725}	Primase
NSP9	N^{3698}-Q^{3805}	N^{3711}-Q^{3818}	N^{3726}-Q^{3833}	
NSP10	A^{3806}-Q^{3940}	A^{3819}-Q^{3953}	A^{3834}-Q^{3968}	
NSP11	T^{3941}-L^{3958}	T^{3954}-L^{3971}	T^{3969}-L^{3986}	Short peptide at the end of ORF1a
NSP12	T^{3941}-Q^{4867}	T^{3959}-Q^{4880}	T^{3964}-Q^{4895}	RdRp
NSP13	S^{4868}-Q^{5464}	A^{4881}-Q^{5477}	A^{4896}-Q^{5492}	HEL, NTPase
NSP14	A^{5465}-Q^{5982}	A^{5478}-Q^{5995}	A^{5493}-Q^{6010}	ExoN, NMT
NSP15	S^{5983}-Q^{6321}	S^{5996}-Q^{6330}	S^{6011}-Q^{6348}	NendoU
NSP16	S^{6322}-V^{6623}	S^{6331}-K^{6631}	S^{6349}-V^{6646}	OMT

[a] ADRP ADP-ribose 1-phosphatase, PL2pro papain-like protease 2, 3CLpro coronavirus NSP5 protease, Hel helicase, NTPase nucleoside triphosphatase, ExoN exoribonuclease, NMT N7 methyltransferase, NendoU endoribonuclease, OMT 2′ O-methyltransferase.

transcription regulatory sequences (TRS-L), and the putative body TRSs, representing signals for the discontinuous transcription of subgenomic mRNAs (sgmRNAs), have been identified in the three genomes (Table 1). The TRS-L and TRSs preceded the codon start of all ORFs in BatCoV-Ita3 and suggested the synthesis of 6 monocistronic subgenomic mRNAs. The lack of TRSs before the ORF3 gene codon start in BatCoV-Ita4 and BatCoV-Ita5 suggests the synthesis of 4 monocistronic and 1 polycistronic subgenomic mRNAs. The differences at the nucleotide level were also confirmed at the amino acid level. BatCoV-Ita4 and BatCoV-Ita5 showed high similarities (< 97%) and high differences with BatCoV-Ita3 in the spike, ORF3 and nucleocapsid proteins. The ICTV has established the 90% amino acid sequence identity of the seven concatenated domains within the ORF1ab as the threshold value to assign two strains to the same species: NSP3 (ADRP), NSP5 (3CLpro), NSP12 (RdRp), NSP13 (Hel, NTPase), NSP14 (ExoN, NMT), NSP15 (NendoU), and NSP16 (OMT). To classify the Italian strains into known coronavirus species, the ORF1ab concatenated domains were compared with the 11 Alpha-CoV species: *Miniopterus bat coronavirus 1, Bat coronavirus CDPHE15, Miniopterus bat coronavirus HKU8, Rhinolophus bat coronavirus HKU2, Bat coronavirus HKU10, Scotophilus bat coronavirus 512, PEDV, HuCoV-229E, HuCoV-NL63, AlphaCoVs1,* and some strains that to date are not assigned. BatCoV-Ita3 concatenated domains showed sequence identities < 83.8% with all the Alpha-CoV strains. BatCoV-Ita4 and BatCoV-Ita5 shared 99.3% identity and had < 79.1% with all other Alpha-CoVs, suggesting that the classification of the Italian strains should be into two novel Alpha-CoVs species.

3.3. Phylogenetic analyses

The RDP4 recombination detection methods, applied to the dataset of CoVs complete genomes to detect the occurrence of recombination, supported the absence of recombination between the Italian strains and the Alpha-CoVs strains (P values > 0.05). As shown in the ML tree built with complete genomes (Fig. 1), the Italian strains clustered with the Chinese strain BtNv-AlphaCoV/SC2013 (KJ473809) out of the monophyletic clade formed by the complete genomes of the HKU-8, 1 A, 1B and HKU10 species. The former cluster is divided into two sub-clusters: one sub-cluster represented by BatCoV-Ita3 and BtNv-AlphaCoV/SC2013, sharing 75% nt. id., and the other sub-cluster represented by BatCoV-Ita4 and BatCoV-Ita5, sharing 71% nt. id. with the Chinese strain.

The Italian strains showed approximately 62% nt. id. with the strains classified into the HKU10 species (Hipposideros bat coronavirus HKU10 isolate LSH5A, Rousettus bat coronavirus HKU10 isolate 183 A) and < 60% with all other AlphaCoV strains (Supplementary Fig. 1, Supplementary Fig. 2). Additionally, at the amino acid level, the Italian strains showed the highest identities with the Chinese BtNv-AlphaCoV/SC2013 strain with respect to the other Alpha-CoVs. BatCoV-Ita3 showed high identities in all predicted proteins excepting in the ORF3 and N proteins. BatCoV-Ita4 and BatCoV-Ita5 showed lower identities with respect to BatCoV-Ita3, which showed high identities in the orf1ab polyprotein and M proteins (> 75%) and low identities in the other predicted proteins.

The Bayesian trees, built using the predicted protein sequences of E, M and N, confirmed the clustering of the Italian strains with the Chinese strain BtNv-AlphaCoV/SC2013 (data not shown). The tree built with S protein sequences showed a uniquely supported clade, containing the Italian strains, the BtNv-AlphaCoV/SC2013, HKU10, 1 A, 1B, and HKU8 species and the unclassified strain BtMr-AlphaCoV/SAX2011, suggesting correlation only between those bat species (Supplementary Fig. 3).

The Italian bat strains showed low identities with the HuCoV-229E (< 49%) and HuCoV-NL63 (< 45%) strains at the spike protein level and had < 45% identity with HuCoV-229E and < 35% with HuCoV-NL63 at the Receptor Binding Domain (RBD) level. The prediction structure of the spike protein showed a type I membrane glycoprotein divided into two subunits (S1 and S2), as other Alpha-CoVs Spike proteins with most of the protein exposed on the outside of the virus and two transmembrane domains located at the C terminus. However, the Italian strains did not exhibit significant or supported similarities to the known secondary structure receptor-binding domains (HuCoV-229E, -NL63) using the online tool Phyre2 or the Swiss model due to their high divergences (data not shown).

To investigate the correlation among Alpha-CoV strains previously detected worldwide, a phylogenetic tree of the partial RdRp gene was built (Supplementary Fig. 4). The ML showed that strains detected in the same continent shared > 89% nt. id. and were correlated, forming monophyletic clusters while sequences from a different cluster showed a nt. id. < 85%. Most of the Alpha-CoVs species were detected in the same continent as the 1 A, 1B, Bat-CoV 512, HKU2, and HKU8 species in Asia or the CDPHE15 species in North America. Bat coronaviruses related to human HuCoV-229E were retrieved in Africa and the coronaviruses related to HuCoV-NL63, in Africa and America. The HKU10 CoV strains showed sequences similar to those detected in Asia and Europe. Some strains formed a cluster outside of those classified into known Alpha-CoV species.

The Italian strains formed two clusters with the Chinese strain BtNv-AlphaCoV/SC2013 and some European strains. At the RdRp partial gene level, the BtNv-AlphaCoV/SC2013 strain showed 83% nucleotide identity with BatCoV-Ita3, 83.8% with BatCoV-Ita4 and 82.8% with BatCoV-Ita5. The first cluster is formed by BatCoV-Ita4 and BatCoV-Ita5, one Italian strain and one Spanish strain (P.kuh/Iprima/Spain/2007, HQ184058), collected from the bat *Pipistrellus kuhlii* species in the Southwest Piedmont region in Northern Italy (Pkuh605, KY780383) and in Spain in 2014 and 2007. These strains shared > 96.7% nt. id. The second cluster contains the BatCoV-Ita3 with one Italian strain collected in the centre of the Piedmont region in Northern Italy (Ppip1015C, KY780385), and a French strain (KT345294, Pip1_Cr_FR_2014), both collected in 2014. Those strains formed a monophyletic clade with two European strains, detected in Bulgaria (GU190239, BNM98-30/BGR/2008) and Spain (HQ184057, M.myo/I/Spain/2007), and two strains from South Africa (KF843855, BtCoV/GrNC1/Neo; KF843862, BtCoV/GrNC8/Neo) from the *Nyctalus leisleri*, *Myotis myotis* and *Neoromicia capensis* species, sharing with BatCoV-Ita3 approximately 83% nt. id.

4. Discussion

In this study, three Alpha-CoV strains from the *Pipistrellus kuhlii* bat species were fully sequenced. The *P. kuhlii* species is one of the most frequently described bat species in Italy that forages in urban and agricultural areas (Russo and Jones, 2003; Ancillotto et al., 2016).

To fully characterize the three Alpha-CoV strains, the NGS method previously described by Moreno et al. (2017) was applied successfully, obtaining the Alpha-CoV complete genome sequences with high coverage rates. However, the lack of European Alpha-CoV complete genomes make difficult to conduct a comprehensive genetic and phylogenetic analysis. The analysis on the full BatCoV-Ita sequences showed similarities to the Alpha-coronavirus genera and genome organization with 6 open reading frames (ORFs) and the 5′ and 3′ non-translated sequences. The phylogenetic analysis using the complete genomes showed correlation but with a low nucleotide identity with a Chinese strain detected in 2013 in the *Nyctalus velutinus* species. The phylogenetic analysis on amino acidic sequences also confirmed the correlation with the Chinese strain and supported the hypothesis that bat strains of *Miniopterus bat coronavirus 1, Miniopterus bat coronavirus HKU8,* and *Bat coronavirus HKU10* species and some unclassified strains may share a common spike ancestor. However, the analysis of the protein structure was hampered by the lack of similar spike protein structure. Indeed, due to the high genetic divergences with human strains it was impossible to predict the spike structure and the affinity with the human receptor.

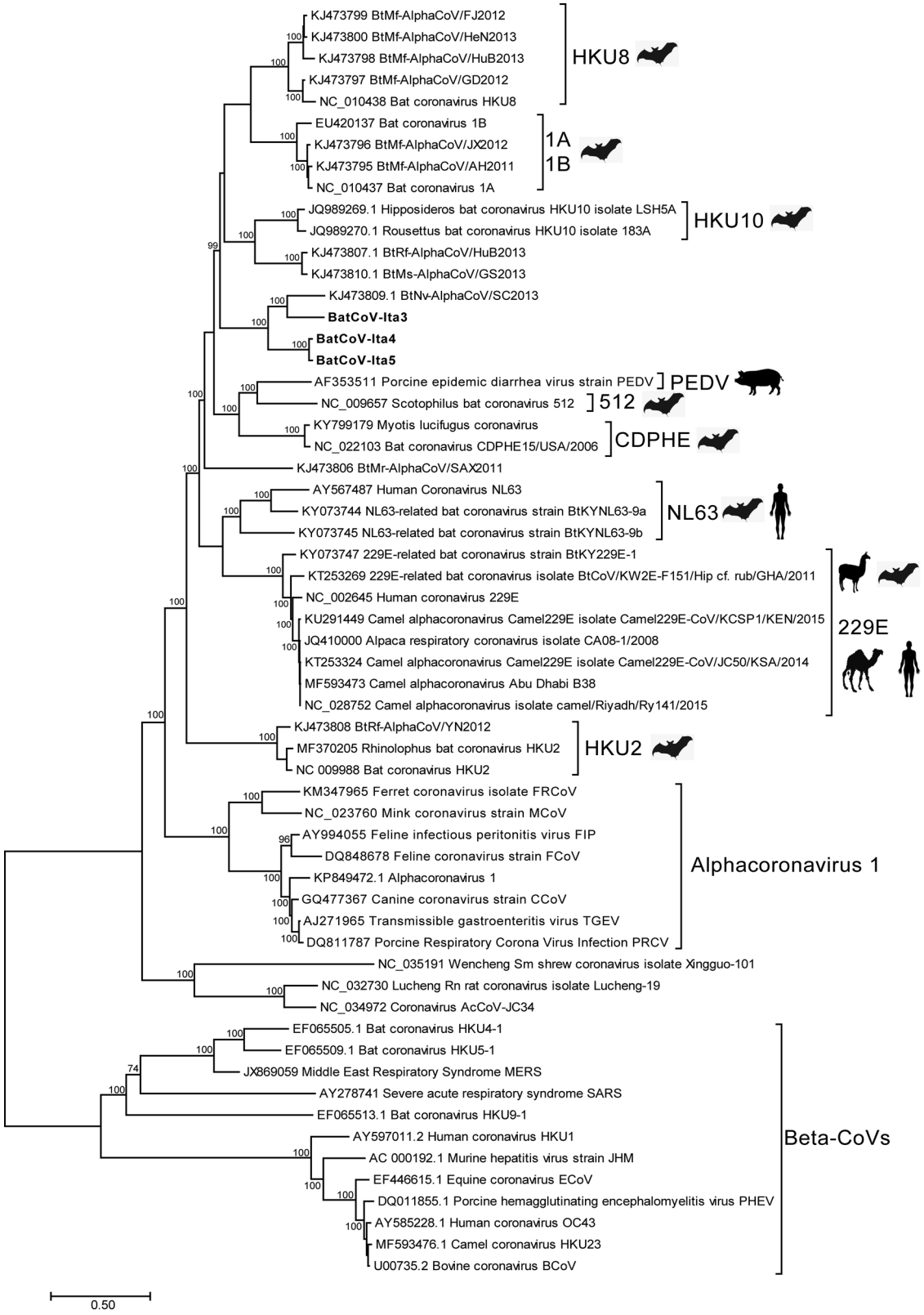


Fig. 1. Maximum phylogenetic tree based on 47 Alpha-CoVs and 12 Beta-CoVs complete genomes. The tree was inferred under the GTR + G + I substitution model and 1000 bootstrap resampling process replications showing values > 70. BatCoV-Ita3, BatCoV-Ita4 and BatCoV-Ita5 are reported in bold and the sequences can be retrieved under accession numbers (MH938448, MH938449, MH938450).

The ICTV has established that viruses sharing more than 90% amino acid sequence identity in the conserved concatenated domains of the orf1ab polyprotein can be assigned to the same CoV species (https://talk.ictvonline.org/taxonomy/). The ICTV demarcation criteria for genera and species allowed us to classify the BatCoV-Ita into two novel Alpha-CoVs species. Our results support previous findings about the high heterogeneity of CoVs hosted by bats and support the idea that novel species may be found in the future with increasing surveillance.

Several studies described the presence of Alpha-CoV and Beta-CoVs in bats worldwide (Falcon et al., 2011; Gouilh et al., 2011; August et al., 2012; Goffard et al., 2015; Asano et al., 2016; Fischer et al., 2016; Goes et al., 2016; Subudhi et al., 2017; Ar Gouilh et al., 2018; Geldenhuys et al., 2018). However, most of these studies reported phylogenetic analysis on short sequences within the RdRp region, establishing the correlation with other CoV strains but not the assignment to CoV species as established by ICTV.

In Europe, CoV strains were detected in samples from more than 20 different bat species. In Italy, a large variety of CoV strains were detected in *Myotis nattereri, Myotis daubentonii, Myotis myotis, Rhinolophus hipposideros, Hypsugo savii, Pipistrellus kuhlii, Pipistrellus pipistrellus, Nyctalus noctula, Epseticus serotinus, Myotis blythii, Myotis oxygnathus,* and *Plecotus auritus* (Lelli et al., 2013; De Benedictis et al., 2014; Rizzo et al., 2017).

The phylogenetic analysis on the RdRp region showed the correlation of the Alpha-CoV strains detected in the same continent. Interestingly, within each geographic area most of the strains hosted by the same bat genera cluster together, confirming the CoVs -host coevolution. BatCoV-Ita4 and BatCoV-Ita5 strains showed high nucleotide identity with one Italian strain and one Spanish strain detected in 2014 and 2007, respectively, from the bat *Pipistrellus kuhlii* species (HQ184058.1; KY780383.1) (Falcon et al., 2011; Rizzo et al., 2017). The BatCoV-Ita3 result correlated with one Italian and one French strain (KY780385.1; KT345294.1), both collected in 2014, from the bat *Pipistrellus Pipistrellus* species (Goffard et al., 2015; Rizzo et al., 2017). The high identities at the RdRp gene level and the clustering of the European strains with the Italian strains suggest that the two novel Alpha-CoV species detected in this study may infect at least two bat species of the *Pipistrellus* genera (*Pipistrellus kuhlii* and *Pipistrellus Pipistrellus*) from different European countries.

In contrast, some geographical clusters were represented by strains detected in different bat genera, attesting to the capability of the CoVs interspecies jumping that may occur when different species of bats share same roost (Leopardi et al., 2018).

Bat behaviour, including flying long distances, living in large colonies, having social interactions, and cohabitating with different bat species, favour the interspecies or intraspecies transmission of viruses (Calisher et al., 2006). During the last fifteen years, two Beta-CoVs, SARS-CoV and MERS-CoV, have jumped from bats to a mammalian intermediate host to humans (Field, 2009; Omrani et al., 2015). In addition, strains related to human Alpha-CoVs (HuCoV-229E, HuCoV-NL63) have been detected in bats, indicating the importance of the bat as a CoV reservoir (Pfefferle et al., 2009; Corman et al., 2015; Tao et al., 2017; Waruhiu et al., 2017). In this study, we characterized two Alpha-CoVs from Italian bats divergent from human CoVs strains and two new Alpha-CoV species. In addition, the RdRp phylogenetic tree showed that the strains here described were not related to the Alpha-CoV species established so far. This result highlights that the heterogeneity of CoVs in the bat may be higher than what is known to date. Indeed, to better understand the CoV species circulating in bats, their evolution and our understanding of the mechanisms important to cross the species barrier, it is important to have long-term vigilance followed by the complete genome characterization.

Funding

This study was partially funded by the Italian Ministry of Health under the Research projects: GR-2011-02350591 PGR2011001- “An epizootiological survey of bats as reservoirs of emerging zoonotic viruses in Italy: implications for public health and biological conservation”; PE-2011-02351681 PRF2011301 “Emerging respiratory viruses: monitoring of coronavirus infections at the human-animal interface”. The funding body was not involved into the design of the study, and collection, analysis, and interpretation of data in the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2018.11.007.

References

Ancillotto, L., Santini, L., Ranc, N., Maiorano, L., Russo, D., 2016. Extraordinary range expansion in a common bat: the potential roles of climate change and urbanization. Naturwissenschaften 103 (3–4), 15.
Annan, A., Baldwin, H.J., Corman, V.M., et al., 2013. Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. Emerg. Infect. Dis. 19 (3), 456–459.
Apweiler, R., Attwood, T.K., Bairoch, A., et al., 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res. 29 (1), 37–40.
Ar Gouilh, M., Puechmaille, S.J., Diancourt, L., et al., 2018. SARS-CoV related Betacoronavirus and diverse Alphacoronavirus members found in western old-world. Virology 517, 88–97.
Asano, K.M., Hora, A.S., Scheffer, K.C., et al., 2016. Alphacoronavirus in urban Molossidae and Phyllostomidae bats. Brazil. Virol. J. 13, 110.
August, T.A., Mathews, F., Nunn, M.A., 2012. Alphacoronavirus detected in bats in the United Kingdom. Vector Borne Zoonotic Dis. 12 (6), 530–533.
Bateman, A., Birney, E., Cerruti, L., et al., 2002. The Pfam protein families database. Nucleic Acids Res. 30 (1), 276–280.
Calisher, C.H., Childs, J.E., Field, H.E., et al., 2006. Bats: important reservoir hosts of emerging viruses. Clin. Microbiol. Rev. 19 (3), 531–545.
Corman, V.M., Baldwin, H.J., Tateno, A.F., et al., 2015. Evidence for an ancestral association of human coronavirus 229E with bats. J. Virol. 89 (23), 11858–11870.
Corman, V.M., Ithete, N.L., Richards, L.R., et al., 2014. Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. J. Virol. 88 (19), 11297–11303.
De Benedictis, P., Marciano, S., Scaravelli, D., et al., 2014. Alpha and lineage C BetaCoV infections in Italian bats. Virus Genes 48 (2), 366–371.
Djikeng, A., Halpin, R., Kuzmickas, R., et al., 2008. Viral genome sequencing by random priming methods. BMC Genomics 9, 5.
Drexler, J.F., Corman, V.M., Drosten, C., 2014. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. Antiviral Res. 101, 45–56.
Falcon, A., Vazquez-Moron, S., Casas, I., et al., 2011. Detection of Alpha and Betacoronaviruses in multiple Iberian bat species. Arch. Virol. 156 (10), 1883–1890.
Field, H.E., 2009. Bats and emerging zoonoses: henipaviruses and SARS. Zoonoses Public Health 56 (6-7), 278–284.
Fischer, K., Zeus, V., Kwasnitschka, L., et al., 2016. Insectivorous bats carry host specific astroviruses and coronaviruses across different regions in Germany. Infect. Genet. Evol. 37, 108–116.
Geldenhuys, M., Mortlock, M., Weyer, J., et al., 2018. A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in Neoromicia bats within South Africa. PLoS One 13 (3), e0194527.
Gloza-Rausch, F., Ipsen, A., Seebens, A., et al., 2008. Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. Emerg. Infect. Dis. 14 (4), 626–631.
Goes, L.G.B., Campos, A.C.A., Carvalho, C., et al., 2016. Genetic diversity of bats coronaviruses in the Atlantic Forest hotspot biome. Brazil. Infect. Genet. Evol. 44, 510–513.
Goffard, A., Demanche, C., Arthur, L., et al., 2015. Alphacoronaviruses detected in french bats are phylogeographically linked to coronaviruses of european bats. Viruses 7 (12), 6279–6290.
Gouilh, M.A., Puechmaille, S.J., Gonzalez, J.P., et al., 2011. SARS-Coronavirus ancestor’s foot-prints in South-East Asian bat colonies and the refuge theory. Infect. Genet. Evol. 11 (7), 1690–1702.
Holmes, E.C., 2009. The evolution and emergence of RNA viruses. In: Harvey, P.H., May, R.M. (Eds.), Oxford Series in Ecology and Evolution. Oxford University Press Inc., New York.
Hu, B., Ge, X., Wang, L.F., et al., 2015. Bat origin of human coronaviruses. Virol. J. 12, 221.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.
Huynh, J., Li, S., Yount, B., et al., 2012. Evidence supporting a zoonotic origin of human coronavirus strain NL63. J. Virol. 86 (23), 12816–12825.
Ithete, N.L., Stoffberg, S., Corman, V.M., et al., 2013. Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. Emerg. Infect. Dis. 19 (10), 1697–1699.
Kemenesi, G., Dallos, B., Gorfol, T., et al., 2014. Molecular survey of RNA viruses in Hungarian bats: discovering novel astroviruses, coronaviruses, and caliciviruses. Vector Borne Zoonotic Dis. 14 (12), 846–855.
Kiemer, L., Lund, O., Brunak, S., et al., 2004. Coronavirus 3CLpro proteinase cleavage sites: possible relevance to SARS virus pathology. BMC Bioinformatics 5, 72.
Lai, M.M., 1992. RNA recombination in animal and plant viruses. Microbiol. Rev. 56 (1), 61–79.
Lau, S.K., Li, K.S., Tsang, A.K., et al., 2013. Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. J. Virol. 87 (15), 8638–8650.
Lelli, D., Papetti, A., Sabelli, C., et al., 2013. Detection of coronaviruses in bats of various species in Italy. Viruses 5 (11), 2679–2689.

Leopardi, S., Holmes, E.C., Gastaldelli, M., et al., 2018. Interplay between co-divergence and cross-species transmission in the evolutionary history of bat coronaviruses. Infect. Genet. Evol. 58, 279–289.
Martin, D.P., Murrell, B., Golden, M., et al., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1 (1) vev003.
Monchatre-Leroy, E., Boue, F., Boucher, J.M., et al., 2017. Identification of Alpha and Beta Coronavirus in Wildlife Species in France: Bats, Rodents, Rabbits, and Hedgehogs. Viruses 9 (12).
Moreno, A., Lelli, D., de Sabato, L., et al., 2017. Detection and full genome characterization of two Beta CoV viruses related to Middle East respiratory syndrome from bats in Italy. Virol. J. 14 (1), 239.
Omrani, A.S., Al-Tawfiq, J.A., Memish, Z.A., 2015. Middle East respiratory syndrome coronavirus (MERS-CoV): animal to human interaction. Pathog. Glob. Health 109 (8), 354–362.
Pauly, M., Pir, J.B., Loesch, C., et al., 2017. Novel alphacoronaviruses and paramyxoviruses cocirculate with type 1 and severe acute respiratory system (SARS)-Related betacoronaviruses in synanthropic bats of Luxembourg. Appl. Environ. Microbiol. 83 (18).
Pfefferle, S., Oppong, S., Drexler, J.F., et al., 2009. Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats. Ghana. Emerg. Infect. Dis. 15 (9), 1377–1384.
Reusken, C.B., Lina, P.H., Pielaat, A., et al., 2010. Circulation of group 2 coronaviruses in a bat species common to urban areas in Western Europe. Vector Borne Zoonotic Dis. 10 (8), 785–791.
Rizzo, F., Edenborough, K.M., Toffoli, R., et al., 2017. Coronavirus and paramyxovirus in bats from Northwest Italy. BMC Vet. Res. 13 (1), 396.
Russo, D., Jones, G., 2003. Use of foraging habitats by bats in a Mediterranean area determined by acoustic surveys: conservation implications. Ecography. 26, 197–209.
Shi, Z., 2013. Emerging infectious diseases associated with bat viruses. Science China. Life Sci. 56 (8), 678–682.
Simmonds, P., 2012. SSE: a nucleotide and amino acid sequence analysis platform. BMC Res. Notes 5, 50.
Smith, I., Wang, L.F., 2013. Bats and their virome: an important source of emerging viruses capable of infecting humans. Curr. Opin. Virol. 3 (1), 84–91.
Subudhi, S., Rapin, N., Bollinger, T.K., et al., 2017. A persistently infecting coronavirus in hibernating Myotis lucifugus, the North American little brown bat. J. Gen. Virol. 98 (9), 2297–2309.
Tao, Y., Shi, M., Chommanard, C., et al., 2017. Surveillance of bat coronaviruses in Kenya identifies relatives of human coronaviruses NL63 and 229E and their recombination history. J. Virol. 91 (5).
Waruhiu, C., Ommeh, S., Obanda, V., et al., 2017. Molecular detection of viruses in Kenyan bats and discovery of novel astroviruses, caliciviruses and rotaviruses. Virol. Sin. 32 (2), 101–114.
Weiss, S.R., Leibowitz, J.L., 2011. Coronavirus pathogenesis. Adv. Virus Res. 81, 85–164.
Wolfe, N.D., Dunavan, C.P., Diamond, J., 2007. Origins of major human infectious diseases. Nature 447 (7142), 279–283.
Woo, P.C., Lau, S.K., Huang, Y., et al., 2009a. Coronavirus diversity, phylogeny and interspecies jumping. Exp. Biol. Med. (Maywood) 234 (10), 1117–1127.
Woo, P.C., Lau, S.K., Lam, C.S., et al., 2009b. Comparative analysis of complete genome sequences of three avian coronaviruses reveals a novel group 3c coronavirus. J. Virol. 83 (2), 908–917.
Woo, P.C., Lau, S.K., Lam, C.S., et al., 2012. Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of Alphacoronavirus and Betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J. Virol. 86 (7), 3995–4008.
Woo, P.C., Lau, S.K., Yuen, K.Y., 2009c. Clinical features and molecular epidemiology of coronavirus-HKU1-associated community-acquired pneumonia. Hong Kong Med. J. = Xianggang yi xue za zhi 15 (Suppl 9), 46–47.

SUPPLEMENTARY MATERIAL

Supplementary Table 1 Pairwise amino acid sequence identity between BatCoV-Ita3 and Alpha-CoV strains

Alphacoronavirus[a]	BatCoV-Ita3					
	ORF1ab	S	ORF3	E	M	N
BatCoV-Ita4/5	75.5-75.7	64.1-64.1	55.9-56.3	70.7	88.1	60.1-60.5
BtNv-AlphaCoV/SC2013	82	76.6	64.5	90.7	93.8	62.6
HKU8	64.1	47.3	46.2	52.7	72.2	43.7
Mi-BatCoV 1A	64.1	44.9	38.2	55.4	70.9	47.3
Mi-BatCoV 1B	64	44.3	39.1	56.8	70	47.4
Ro-BatCoV HKU10 183A	64.6	48	42.1	56	71.4	48.4
Hi-BatCoV HKU10 LSH5a	64.5	49.1	42.1	56	71.7	48.2
PEDV	64.2	43.8	36.4	52.7	65.9	41.4
Sc-BatCoV 512	63.5	43	38.3	45.9	67	47.8
CDPHE15/USA/2006	63.5	44.3	36.8	43.8	63.3	42.1
HCoV NL63	60.8	44.8	41.4	49.3	63.3	48.1
HCoV 229E	60.4	49.2	30.1	44	58.8	40.3
HKU2	60.6	26.9	33.5	41.3	64.3	48.8

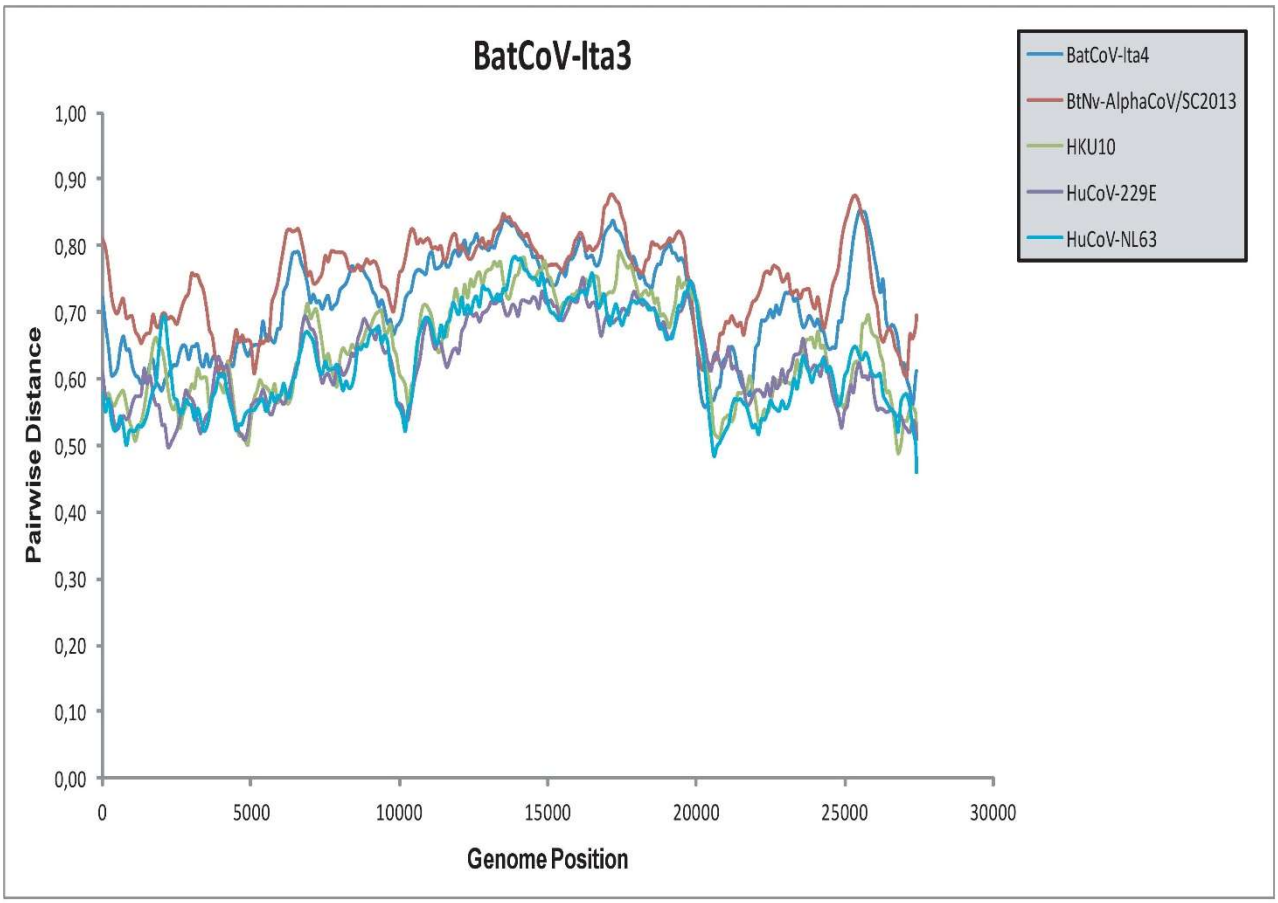
[a]Alphacoronavirus accession number: BtNv-AlphaCoV/SC2013 KJ473809, HKU8 NC_010438, BatCoV 1A NC_010437, BatCoV 1B EU420137, HKU10 LSH5A JQ989269, HKU10 183A JQ989270, PEDV Porcine epidemic diarrhea virus strain AF353511, BatCoV 512 NC_009657, Bat coronavirus CDPHE, Hu-CoV NL63 AY567487, Hu-CoV 229E NC_002645, HKU2 MF370205

Supplementary Table 2 Pairwise amino acid sequence identity between BatCoV-Ita4 and BatCoV-Ita5 with Alpha-CoV strains

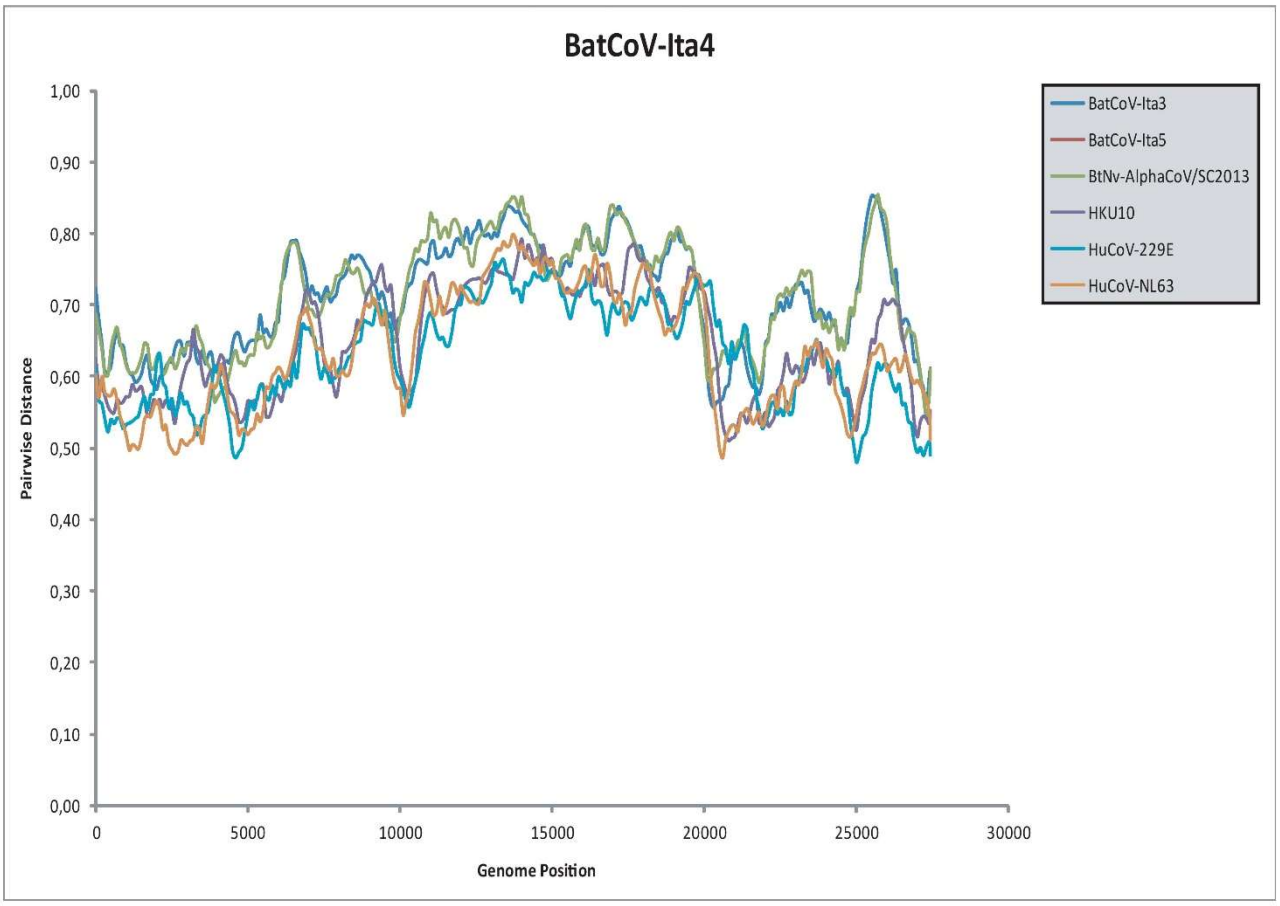
Alphacoronavirus[a]	BatCoV-Ita4/5					
	ORF1ab	S	ORF3	E	M	N
BatCoV-Ita3	75.5-75.7	64.1-64.1	55.9-56.3	70.7	88.1	60.1-60.5
BtNv-AlphaCoV/SC2013	75.9	65.6	60-61.5	68	84.7	64
HKU8	64	48-48.2	40.7-41.1	52.7	67.2	47-47.2
BatCoV 1A	64.2-64.3	44.8-45	36.4-36.8	55.4	66.4	52.1-52.5
BatCoV 1B	64.4	45.3-45.4	35.9-36.4	56.8	66.8	51.6-51.7
HKU10 183A	64.9	47.1-47.3	40.7-41.1	58.7	72	52.5
HKU10 LSH5a	64.7	48-48.1	40.7-41.1	58.7	71.7	52.5
PEDV	63.4-64	45.3-45.5	37	59.5	64.1	44.6-44.8
BatCoV 512	63.3	43.8-44.4	35.6-36.1	51.3	63.9	48.1-48.5
BatCoV CDPHE	63.3-63.4	44.3-44.6	36.4-36.9	60.3	61.9	42.4-42.9
HuCoV-NL63	61.2	45.1	33-34	50.7	64.1	45.9-46
HuCoV-229E	60.2-60.3	48.3-48.7	28-28.8	50.7	56.4	40.6
Rh-BatCoV HKU2	60.5	28.2	30.9-31.8	52	60.2	49.7-49.9

[a]Alphacoronavirus accession number: BtNv-AlphaCoV/SC2013 KJ473809, HKU8 NC_010438, BatCoV 1A NC_010437, BatCoV 1B EU420137, HKU10 LSH5A JQ989269, HKU10 183A JQ989270, PEDV Porcine epidemic diarrhea virus strain AF353511, BatCoV 512 NC_009657, Bat coronavirus CDPHE, Hu-CoV NL63 AY567487, Hu-CoV 229E NC_002645, HKU2 MF370205

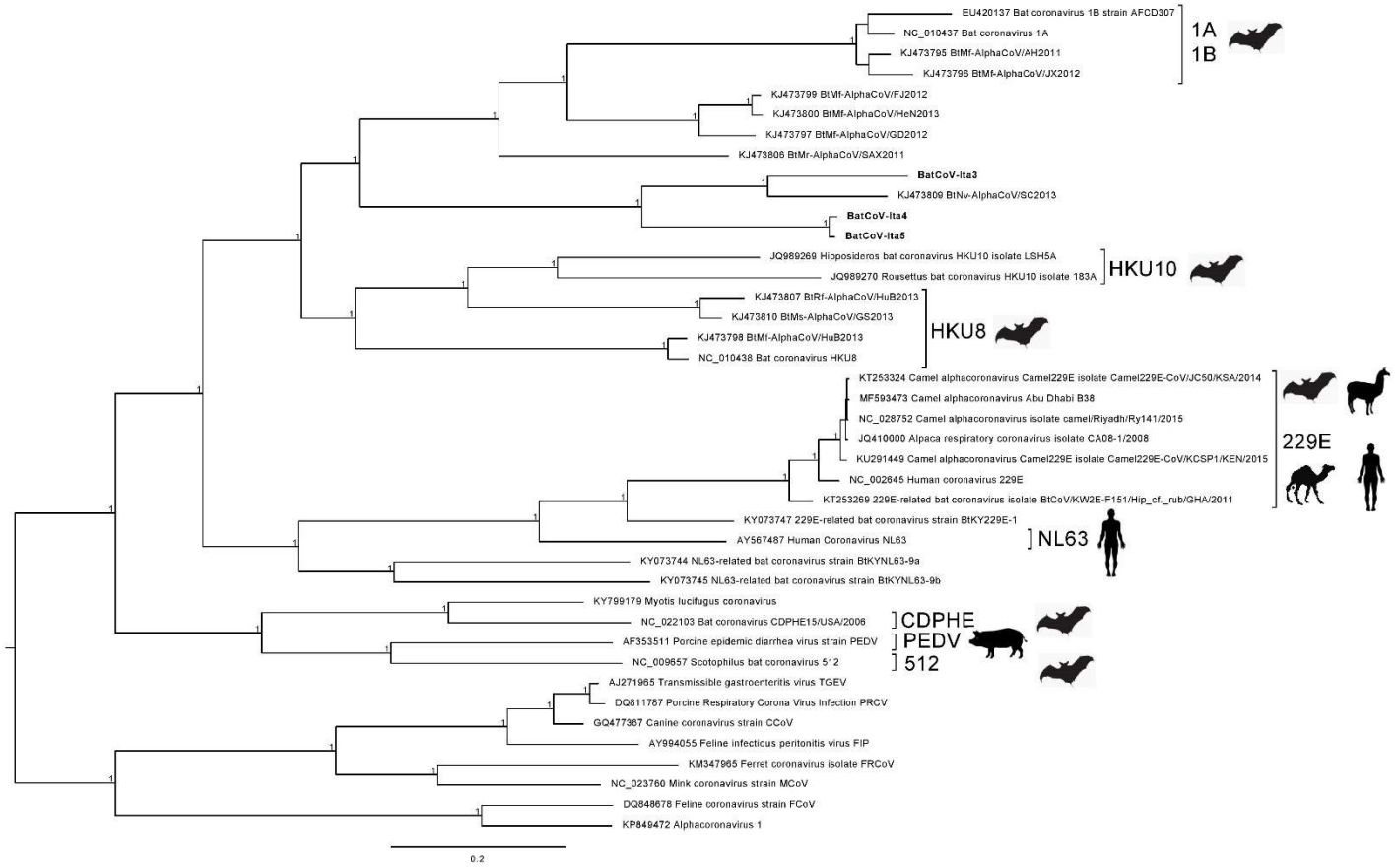
Supplementary Fig. 1 Sequence identity between BatCoV-Ita3 with other AlphaCoV strains: BatCoV-Ita4, BtNv-AlphaCoV/SC2013, HKU10, HuCoV-NL63, HuCoV-229E. Similarity plots were generated using SSE version 1.2 using a sliding window of 600 and a step size of 100 nucleotides



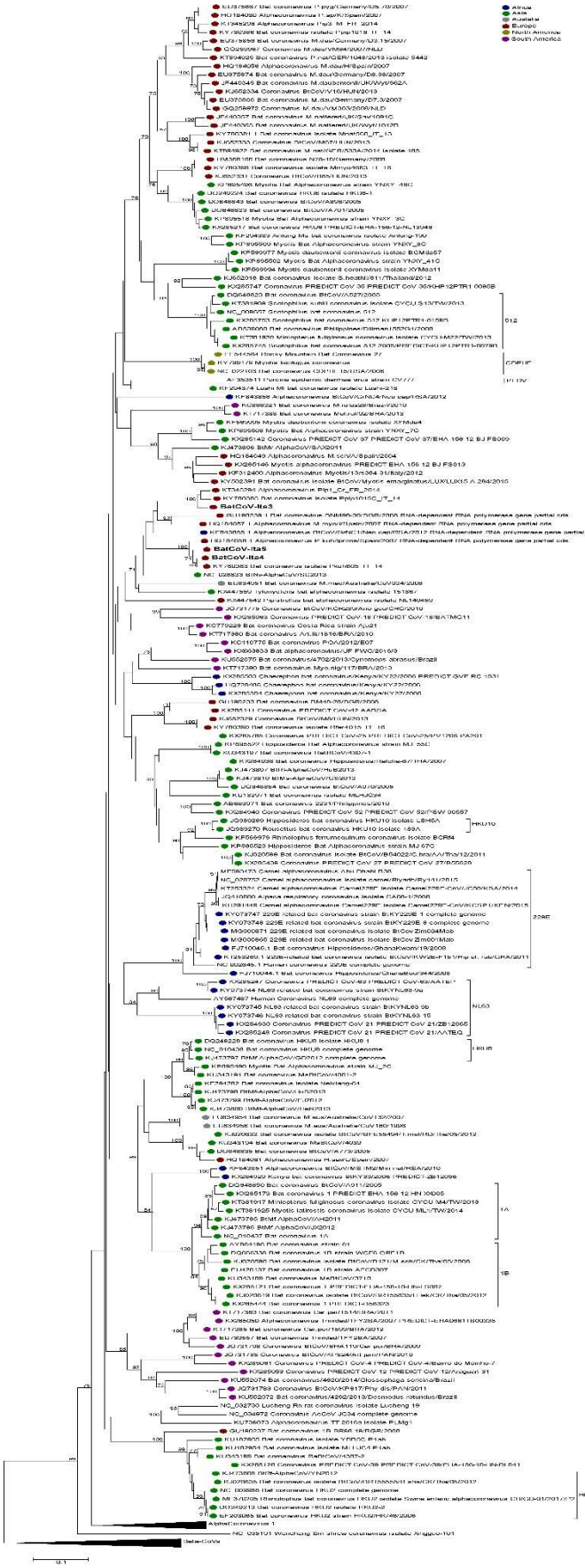
Supplementary Fig. 2 Sequence identity between BatCoV-Ita4 with other AlphaCoV strains: BatCoV-Ita3, BatCoV-Ita5, BtNv-AlphaCoV/SC2013, HKU10, HuCoV-NL63, HuCoV-229E. Similarity plots were generated using SSE version 1.2 using a sliding window of 600 and a step size of 100 nucleotides



Supplementary Fig. 3 Bayesian phylogenetic tree of spike protein sequences excluding the divergent HKU2 sequences strains. BatCoV-Ita3, BatCoV-Ita4, BatCoV-Ita5 are reported in bold



Supplementary Fig. 4 Maximum phylogenetic tree based on 191 Alpha-CoV and 12 Beta-CoV RdRp sequences as outgroup. The tree was inferred under the GTR + G + I substitution model and 1000 bootstrap resampling process replications showing values > 70. BatCoV-Ita3, BatCoV-Ita4 and BatCoV-Ita5 are reported in bold.



CHAPTER 9

Multiplex amplification method for HEV-3 and MERS-CoV complete genomes

9. Multiplex amplification method for HEV-3 and MERS-CoV complete genomes

In the previous Chapters, we described the characterization of HEV and MERS-CoV strains, detected in animal reservoir, by phylogenetic analyses on short or complete genome sequences. The genome sequences were obtained applying the NGS method based on random sequencing of all nucleic acids in the samples. However, results were not totally satisfactory. Full genome sequences were not obtained when the method was used to sequence HEV-3 virus from pig feces. Differently, we obtained full genome sequences when the viruses, both HEV-3 and CoVs, were sequenced from wild boar liver or bat carcass/feces. The different results were probably linked to the viral load and the great amount of non-viral sequences in the sample. To overcome the limit of the method and implement the accuracy of the NGS, we used the novel amplicon-based enrichment method for sequencing of targeting viruses. Two specific primer panels were designed for HEV-3 and MERS-CoV, based on a consensus sequence generated by the alignment of the complete genomes and reference sequences available online. The set of primers were used in a multiplexed PCR to amplify the specific target generating overlapping amplicons to cover the whole genome sequence. The developed HEV-3 panel was tested on 5 wild boar HEV strains belonging to different subtypes. Result enabled to obtain full genome sequences from all samples.

The MERS-CoV panel was used to sequence one human MERS-CoV isolated strain and two bat samples, MERS-CoV like positive, described in Chapter 7, with a 76% homology to MERS-CoV. Complete genomes were obtained, indicating the ability to use the panel developed for the MERS-CoV in humans but also from distant MERS-CoV-like strains detected in bats.

The developed NGS method enables to obtain complete genomes by both panels within a single NGS run. The deep sequencing of targeted genomes result in higher read coverage rates reducing the time and the cost of sequencing. The majority of sequencing reads obtained (>80%) correspond to the target viruses. The advantage of this approach is to obtain (the achievement of) thousands of sequences (reads) corresponding to genome of interest, obtaining a consensus sequence with a high accuracy, reducing at least false positive nucleotide changes. Furthermore, sequencing of target virus reduces the complexity of bioinformatic process to obtain the final result.

Although this is just a proof of principle, the obtained results indicated that multiplex amplification NGS method can be a promising detection and characterization assay in veterinary diagnostic and research settings for viruses, belonging to the same taxon classification level (i.e. genotypes or species) that shows limited nucleotide sequence similarities (>75%).

Introduction

Over the past 10 years, tremendous progresses have been made in the field of Next Generation Sequencing (NGS) applications, among which, the discovery of novel viruses and deep characterization of known viruses are promising applications. NGS technology produces a great amount of sequences from the analysed sample that can be assembled into longer ones without use of a reference genome, related sequences or protein databases in order to identify viruses, in a very limited time frame (Mokili et al., 2012). Compared to traditional methods (Rasmussen & Katze, 2016) that cannot be used for screening or surveillance of unknown pathogens, or monitoring virus evolution in real time during an outbreak, NGS enables more accurate and faster characterization of any viruses. The rapid identification and characterization of newly pathogens from humans and animals became a key step to prevent major public and animal health emerging or re-emerging threats. The number of novel viruses described in literature is increasing every year, with papers reporting continuous discovery of novel or variant viruses using NGS (Mokili et al., 2012). The new Ebola virus Bundiubugyo, the pandemic influenza virus H1N1pdm, HIV, human herpes viruses have been quickly detected using NGS applied to clinical human and animal samples (Li et al., 2016). The complete genome sequencing is essential to have adequate information to characterize virus strains and to establish evolutive correlation among them. However, the relative nucleic acids amount of viruses compared to those of bacteria or host is a critical factor to achieve the whole genome sequence of viruses when NGS technique is applied to biological sample (tissue or feces samples for instance). Indeed, nucleic acids of virus may not be present in adequate amount to be detected. To overcome this limit and to increase the sensitivity of NGS, several enrichment methods have been developed including physical method, amplicon sequencing, PCR-generated baits, and solution-based capture techniques (Wylie et al., 2012; Prachayangprecha et al., 2014; Briese et al., 2015; Wylie et al., 2015). In this study, we design a panel of primers pairs in order to amplify the different subtypes of HEV-3 and MERS-CoV viruses to be sequenced by NGS. The primers pairs were designed in order to produce overlapping amplicons covering the entire genomes of the viruses. The performance of the two panels was evaluated respectively on 4 HEV-3 subtypes and on MERS-CoV or MERS-CoV like viruses directly from biological samples of animal reservoir.

Material and methods

Panel design

For design the HEV-3 panel, 71 complete genomes and 11 HEV subtype reference sequences, downloaded from NCBI database (https://www.ncbi.nlm.nih.gov), were aligned by MEGA7 software. For design the MERS-CoV panel, 70 human and camel complete genomes were used. The consensus sequence was generated selecting for each position the most frequent nucleotide, present at least in > 50% of the sequences. The consensus sequences were uploaded into the Ion AmpliSeq Designer online tool to design the primer set sequencing panel (https://ampliseq.com/browse.action). The Ion Ampliseq technology is an amplicon-based enrichment method for targeted next-generation sequencing (NGS) used to obtain overlapping amplicons of 100–400 bp (Thermo Fisher Scientific, Rodano, MI, Italy).

Library preparation and NGS sequencing

Five HEV-3 positive liver samples, one from wild boar hunted in South Italy (Chapter 4) and four from Viterbo province (Chapter 6), classified in different subtypes, were used to test the HEV Ion AmpliSeq panel. The MERS-CoV Ion AmpliSeq panel was tested on RNA extracted from the human MERS-CoV strain growth *in vitro* (KC164505, England1) (Bermingham et al., 2012) and from two bat carcass samples positive for MERS-CoV like strains (MG596802, BatCoV-Ita1; MG596803, BatCoV-Ita2) (Chapter 7). The libraries were prepared using the Ion AmpliSeq Library Kit 2.0 and the Ion AmpliSeq DNA libraries guidelines. The cDNA was quantified by the Qubit dsDNA HS Assay Kit using Qubit 2.0 (Thermo Fisher Scientific). Thirty ng of cDNA obtained by SISPA method (Chapter 3) were subjected to PCR amplification using 5X Ion AmpliSeq HiFi Master Mix with the following steps for HEV: 99º C for 2 minutes, 15 cycles at 99º C for 15’’and 60º C for 8 minutes before holding at 10 °C. For MERS-CoV the same PCR conditions were used amplifying for 4 minutes at 60°C. Primer sequences were cut using the FuPa Reagent and then ligated to the barcodes and to Ion P1 Adaptor (Thermo Fisher Scientific). The libraries were purified with the Agencourt AMPure XP system (Beckman, Milan, Ialy) and quantified on a High Sensitivity DNA Analysis Kit on the Agilent Bioanalyzer (Agilent Technologies, Roma, Italy). The emulsion PCR and the sequencing run were performed using the Ion PGM Hi-Q View OT2 Kit and Ion PGM Hi-Q View Sequencing Kit. The HEV and MERS-CoV libraries were pooled and loaded on Ion 318 v2 Chip and the sequencing run performed by the Ion Personal Machine (PGM) (Thermo Fisher Scientific).

Bioinformatic and phylogenetic analysis

The sequences obtained from NGS run were analysed using the online tool Galaxy Aries and mapping the reads by the Bowtie2 tool as described in Chapter 3. The phylogenetic analysis and p-distance

values calculation were performed using MEGA 7 software. The Maximum Likelihood tree was built using 87 HEV-3 complete genomes and one HEV-4 as outgroup. The analysis was performed under the GTR + G + I substitution model, as suggested by JModelTest 2.0, and a bootstrap resampling process of 1000 replications to assess node support.

Results

The AmpliSeq panels

The HEV-3 AmpliSeq panel resulted in 42 primer pairs for the amplification of 42 overlapping amplicons, from 125 to 275 bp in length and covering 99.82% of the consensus genome sequence uploaded in the software. The MERS-CoV panel included 115 primer pairs to amplify 115 overlapping amplicons, from 125 - 375 bp in length, covering the 99.88% of the genome.

NGS sequencing run

The NGS run produced ~6 million reads. After quality control, 5.9 million reads were used for viral classification. Five million reads were classified as HEV-3 and ~1 million reads into the MERS-CoV species. The viral reads were used for the mapping and genome reconstruction.

HEV data analysis

The cDNA, used in Chapter 5 to obtain the complete genome of the wild boar WB/HEV/NA17ITA15NA strain (HEV-3i) by NGS, was used as control to assess the use of HEV Ion AmpliSeq panel. Using the panel of primers for HEV-3 amplification the NGS run produced 145,696 reads from the WB/HEV/NA17ITA15NA library of which 128,971 (89.4%) were specific for HEV-3 and used to build the complete genome with a mean coverage of 3,694x. Compared to our previous results, obtained on the same sample with random sequencing NGS method (Chapter 5), the number of specific reads was approximatively 100 times higher (128,971 *vs* 1,000) as well as the mean coverage (3,694x *vs* 25x). Comparable results were obtained for the other HEV-3 samples amplified. The complete genome sequences obtained showed a high sequence coverage and high percentage of reads of the target strains (Table 1).

Table 1. Number of HEV specific reads obtained by HEV Ion Ampliseq panel

HEV samples	Total reads	HEV reads	HEV reads %	Mean coverage
WB03VT17	890,216	752000	84.4	19,900
WB110VT17	1,737,702	1713036	98.5	49,000
WB119VT17	924,940	912767	98.6	27,000
WB161VT17	1,421,661	1347205	94.7	35,800

HEV phylogenetic analysis

The four HEV-3 complete genomes were approximately 7.2 kb long and organized into the three typical HEV ORFs (ORF1, ORF2 and ORF3). The HEV-3 complete genomes were classified following the criteria established by (Smith et al., 2016) by phylogenetic analysis and calculation of p-distance values (Table 2; Fig. 1). The phylogenetic analysis showed the clustering of WB03VT17 into the HEV-3f subtype, confirming the previously analysis described in Chapter 6. The Italian strain WB03VT17 showed 90.6% nt. id. with the HEV-3f subtype reference strain (AB369687) and p-distances <86.7 with all the other HEV strain subtype strains. The comparison of the ORF1 (nonstructural proteins) and ORF2 amino acids sequences (capsid protein) confirmed clustering into the HEV-3f, displaying 98% and 98.4% aminoacid identity in the ORF1 and ORF2, respectively with HEV-3f subtype reference strain. Differently, the ORF3 was distant from the HEV3f strains (92%) and more strictly related to the HEV-3l subtype (97.3-98.1%). The WB110VT17 clustered with three HEV-3 strains, recently proposed to be a novel subtype (Wang et al., 2018) the swine strain swX07-E1 detected in Sweden (EU360977) and two human strains detected in Germany in 2013 and 2015 (KJ873911, FR_R strain; KU980235, hGER15-0031). The WB110VT17 strain full genome sequence displayed with the above mentioned strains 89.8%-90% nt. id. and <88.2% nt. id. with the other HEV-3 reference subtype strain sequences. The amino acids sequence comparisons confirmed the clustering obtained by nucleotide analyses with FR_R, swX07-E1 and hGER15-0031 (ORF1 and ORF2 98% aa. id., ORF3 93% aa. id.). Clustering observed in the phylogenetic tree and p-distance calculation suggested the classification of WB110VT17 into the novel proposed subtype represented by the swine strain X07-E1 and the two human strains FR_R and GER15-0031.

The WB119VT17 strain showed the closest nucleotide identity (89.8%) with the wild boar and human strains detected in Italy and in Europe, currently classified in the HEV-3c subtype. The same results

were observed using the amino acids comparison with the HEV-3c reference strain: 97.9% aa. id. In the ORF1, 96.5% in the ORF2 and 93% in the ORF3.

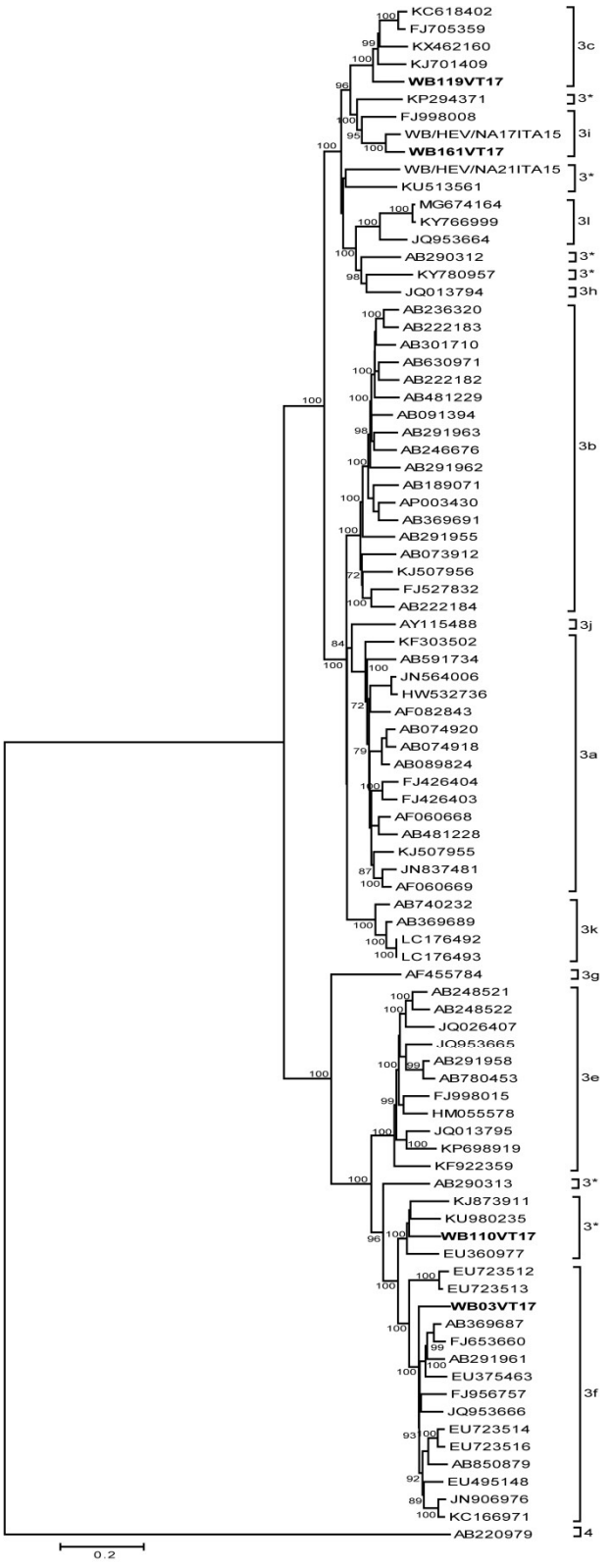
The WB161VT17 strain was the only of the group of the new strains sequenced related to another Italian wild boar strain detected in animal hunted in South Italy. The WB161VT17 showed the highest genetic correlation (93.7% nt. id.) with Italian WB/HEV/NA17ITA15 strains, and with the swine HEV-3i reference strain (FJ998008) (89% nt. id.) confirming their classification in the HEV-3i subtype. The highest amino acids similarities of WB161VT17 were displayed with WB/HEV/NA17ITA15 in ORF1 (98.3%) and 99% in ORF2. Differently, the ORF3 showed 97.3% aminoacid identity with WB110VT17 and only 91.3% with WB/HEV/NA17ITA15.

Table 2. Percent identity values of the Italian wild boar strains with HEV-3 genotype reference sequences for HEV subtypying.

HEV subtype	Accession number	Strain name	Country	WB03VT17	WB110VT17	WB119VT17	WB161VT17
3a	AF082843	Meng	USA	81.5	82.9	84.9	84.3
3b	AP003430	JRA1	Japan	81.4	82.4	84.7	84.9
3c	FJ705359	wbGER27	Germany	81.3	81.9	89.8	85.8
3d*	AF296165	TW12SW	Taiwan	80.6	83.6	83.9	86.5
3e	AB248521	swJ8-5	Japan	84.7	86.4	81.5	81.5
3f	AB369687	E116-YKH98C	Japan	90.8	88.2	81.9	81.6
3g	AF455784	Osh205	Kyrgyzstan	83.1	84.1	81.9	82.4
3h	JQ013794	TR19	France	81.2	82.1	85.9	85.4
3i	FJ998008	BB02	Germany	81.8	82.1	87	89
3j	AY115488	Arkell	Canada	80.8	82.4	84.1	84.1
3k**	-	-	Japan	80.4-80.9	81.9-82.3	84.3-84.4	84.6-85
3l***	-	-	-	80-81.5	81-82.7	84.3-85.6	83.8-85.4
3	AB290312	swMN06-A1288	Mongolia	81.1	82.1	85.5	85.5
3	AB290313	swMN06-C1056	Mongolia	85.6	87	81.6	81.3
3	EU360977	swX07-E1	Sweden	87.4	89.9	81.5	81.7
3	KJ873911	FR_R	Germany	86.8	89.9	81.4	80.9
3	KY780957	SW/16-0282	Switzerland	81.1	81.7	85.2	84.8
3	KU513561	IC2011	Spain	80.9	82	85.8	86
3	KP294371	MWP_2010	Germany	80.9	82.4	86.5	87.3

*Only 304 nt in ORF2 has been reported
** HEV-3k four strains are available, accession numbers LC176493, LC176492, AB369689, AB740232 (strain names: HE-JA16-057, HE-JA16-0578, E088-STM04C and G3-HEV83-2-27, respectively)
*** HEV-3l three strains are available, accession numbers MG674164, KY766999, JQ953664 (strain names: HEV/13RS985-5, SWHEV75BO2012, FR-SHEV3c-like, respectively)

Fig. 1 Maximum likelihood phylogenetic tree reconstruction based on the full-length sequences of 83 HEV-3 strains including the Italian strains WB03VT17, WB110VT17, WB119VT17, WB161VT17 indicated in bold. Bootstraps values>70 are indicated at their respective nodes. Sequences from animals and human strains, belonging to HEV3-a-c and e-k subtypes of genotype 3, have been included in the tree. The HEV-4 strain was used as outgroup.



Coronavirus data analysis

The NGS run on the library prepared by the MERS-CoV panel, used to amplify the human MERS-CoV England1 strain as positive control, produced 750,017 reads of which 749,500 (99.9%) were classified as Coronavirus and used to build the complete genome with a mean coverage rate of 6,150x. The sequences obtained showed 100% nucleotide identity to England1 strain. The MERS-CoV England1 sequence strains showed with the Italian MERS-CoV like bat strains, sequenced in Chapter 6, approximately 76% nucleotide identity. The first NGS method described in Chapter 6 and applied to the Italian bat strains produced 1,991,701 for BatCoV-Ita1 of which 359,827 (18%) used to build the complete genome with 2,000x mean coverage. For BatCoV-Ita2, 1,159,616 reads were retrieved, 150,840 (13.9%) of them were used for the complete genome reconstruction with 870x mean coverage rate. The MERS-CoV panel amplification was applied to the bat MERS-CoV like strains. The sequencing run produced for BatCoV-Ita1 198,813 and for BatCoV-Ita2 283,993 reads. The 78% of the reads obtained (155,463) for BatCoV-Ita1 and 81% of the reads obtained for BatCoV-Ita2 (229,950) were assigned to Coronavirus sequences. The complete genome obtained were 30 kb long with a mean coverage of 766x and 1,600x respectively, showing 100% nucleotide identity with the BatCoV-Ita1 and BatCoV-Ita2 complete genome strains obtained in Chapter 6.

Discussion

An obstacle to obtaining whole RNA virus genome sequence using NGS is the low target loads (titer) in the clinical samples. Our previous study showed recovery of whole genome sequence of HEV-3 strains from liver samples of wild boar (Chapter 6), while only limited sequencing results were obtained when HEV-3 strains are sequenced from pig feces (Chapter 3). The results were not totally satisfactory due to the low accuracy (coverage) of the final genomes built by sequence assembling. To enrich the low amount of viral RNA, in the last part of the doctoral dissertation, a high-throughput method, multiplexed targeted PCR-sequence, using NGS technology was developed for the two viruses under investigation. The performance of the newly developed NGS method was evaluated using 5 HEV-3 strains or MERS-CoV species strains.

Data obtained demonstrated that this approach can provide more accurate sequencing method and confirmed that NGS is a powerful technique for defining virus genome sequences. Despite the low number of samples tested the complete genome of viral strains were retrieved with high coverage rate. Using this method, a significant increase in the number of specific viral read was observed compared to the random amplification and sequencing method performed in Chapter 4 and 6.

The full genome sequences of 4 HEV-3 subtypes were obtained by using the same set of primers. Further studies are needed to define the extend of HEV-3 subtypes that can be sequenced using this set of primers for diagnoses and the ability to discover novel subtypes which may diverge from known

subtype and not be detected by traditional approach. Similar results were obtained with the MERS-CoV panel. The full genome sequence, 30kb, was obtained for the human MERS-CoV strain used as positive control. The complete genomes of MERS-CoV like strains from two bat intestinal contents were also obtained, suggesting the using of the MERS-CoV panel for detection and sequencing of MERS-CoVs and related strains. Furthermore, the reads obtained by NGS belong mainly to the target virus, ~4,7 million out of 5 million for HEV-3 panel and 1 million for MERS-CoV panel out 1,2 for MERS-CoV. The obtained specificity will enable to sequence more samples within a single NGS run with an important save of money and time. The mean coverage obtained in this study was 3,000x and 1,000x for HEV-3 and MERS-CoV like, respectively. The high coverage obtained guaranteed a high accuracy of the resulting consensus sequences. Indeed, almost 30x coverage is considered sufficient for the whole genome sequencing (Sims et al., 2014).

Phylogenetic analysis of the new HEV-3 whole genome sequences obtained confirmed the detection of different HEV-3 subtypes in wild boars. Compared to phylogenetic analysis on short genomic sequences, obtained results enable a more comprehensive classification and leading to a more supported classification when a reference strain is not available or the strains are uncommon. Only for the subtype HEV-3f the short sequences were also be representative. As matter of fact, the WB03VT17 strain was classified as HEV-3f using analysis of either the ORF2 partial sequence or complete genome.

Overall, the sequence analyses revealed 4 different subtypes of HEV-3 in the wild boar analyzed, confirming a wide heterogeneity of the circulating strains. Some subtypes detected, such as HEV-3f and -3c have been already detected in wild boar, pigs and humans in Europe. While the detected HEV-3i, that in this study has been identified in two animals hunted in two separate geographical areas, has not been detected previously in Italy and circulate moderately in Europe.

The comparison of short genomic sequences of the wild boar strains with those available online revealed rather close homology of some strains to isolate from patients in Europe. Unfortunately, only limited sequence information is available on human strains detected in Italy. The WB119VT17 strain (HEV3c) was strictly correlated to an Italian human strain detected in a patient who declared consumption of figatelli (liver sausage) and undercooked wild boar meat (Garbuglia et al., 2015). This is not the first observation of a strict correlation between HEV-3 strains detected in wild boar and in human cases in Italy (Mazzei et al., 2015). The WB110VT17 clustered within a novel HEV-3 subtypes recently proposed (Wang et al., 2018) being closest related to pig strain detected in Sweden and two strains isolated in two patients in German in 2013 and 2015. This result confirms circulation of the same subtype in pig, wild boar and human cases. Wild boar plays an important role as source

of infection of HEV-3, transmissible to humans by consumption of wild boar raw or undercooked meat or sausages that are frequently homemade produced.

In conclusion, the targeted amplification method developed in this study enables to obtain complete genomic sequences in less time and with economic advantage due to the reduce number of reads needed for the successfully complete genome determination. However, the method will be improved to be applied in the future for the diagnosis, surveillance or outbreak monitoring of HEV or MERS-CoV in human and animal reservoir.

References

Bermingham, A., Chand, M.A., Brown, C.S., Aarons, E., Tong, C., Langrish, C., et al., 2012: Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin,* **17,** 20290.

Briese, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O.J., et al., 2015: Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *mBio,* **6**.

Garbuglia, A.R., Alessandrini, A.I., Pavio, N., Tesse, S., Grignolo, S., Viscoli, C., et al., 2015: Male patient with acute hepatitis E in Genoa, Italy: figatelli (pork liver sausage) as probable source of the infection. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases,* **21,** e4-6.

Li, D.C., Li, Z.W., Zhou, Z., Li, Z., Qu, X.Y., Xu, P.S., et al., 2016: Direct next-generation sequencing of virus-human mixed samples without pretreatment is favorable to recover virus genome. *Biology direct,* **11**.

Mazzei, M., Nardini, R., Verin, R., Forzan, M., Poli, A. and Tolari, F., 2015: Serologic and molecular survey for hepatitis E virus in wild boar (Sus scrofa) in Central Italy. *New microbes and new infections,* **7,** 41-47.

Mokili, J.L., Rohwer, F. and Dutilh, B.E., 2012: Metagenomics and future perspectives in virus discovery. *Current opinion in virology,* **2,** 63-77.

Prachayangprecha, S., Schapendonk, C.M.E., Koopmans, M.P., Osterhaus, A.D.M.E., Schurch, A.C., Pas, S.D., et al., 2014: Exploring the Potential of Next-Generation Sequencing in Detection of Respiratory Viruses. *Journal of clinical microbiology,* **52,** 3722-3730.

Rasmussen, A.L. and Katze, M.G., 2016: Genomic Signatures of Emerging Viruses: A New Era of Systems Epidemiology. *Cell host & microbe,* **19,** 611-618.

Sims, D., Sudbery, I., Ilott, N.E., Heger, A. and Ponting, C.P., 2014: Sequencing depth and coverage: key considerations in genomic analyses. *Nature reviews. Genetics,* **15,** 121-132.

Smith, D.B., Simmonds, P., Izopet, J., Oliveira-Filho, E.F., Ulrich, R.G., Johne, R., et al., 2016: Proposed reference sequences for hepatitis E virus subtypes. *The Journal of general virology,* **97,** 537-542.

Wang, B., Harms, D., Papp, C.P., Niendorf, S., Jacobsen, S., Lutgehetmann, M., et al., 2018: Comprehensive Molecular Approach for Characterization of Hepatitis E Virus Genotype 3 Variants. *Journal of clinical microbiology,* **56**.

Wylie, K.M., Mihindukulasuriya, K.A., Sodergren, E., Weinstock, G.M. and Storch, G.A., 2012: Sequence Analysis of the Human Virome in Febrile and Afebrile Children. *PloS one,* **7**.

Wylie, T.N., Wylie, K.M., Herter, B.N. and Storchl, G.A., 2015: Enhanced virome sequencing using targeted sequence capture. *Genome Research,* **25,** 1910-1920.

CHAPTER 10

Discussion and Concluding remarks

10. Discussion and Concluding remarks

Every year new infectious diseases, caused by viral infection, are described (Dong et al., 2008). Over one billion cases of zoonotic diseases are estimated to occur annually. The number of infected people and fatal cases are a major public health concern, together with the economic loss due to impacts on trade, movement of people and economic stability. In the last decades, novel zoonotic pathogens emerged, causing high percentage of infected people worldwide, spreading of the infections and hundreds of billions of dollars' loss (Karesh et al., 2012). Hepatitis E virus and Coronavirus are zoonotic RNA viruses that cause outbreaks and human cases worldwide. Hepatitis E is an acute hepatitis in humans, generally self-limited, that can lead to chronic infections in immunocompromised patients. The zoonotic HEV, classified as genotypes HEV-3 and HEV-4, infects swine, wild boar, and occasionally other animal species (rabbit, deer) without any sign of disease. In Europe, more than 20,000 cases have been reported in the last 10 years with an increasing trend. Foodborne is considered the main route of HEV-3 and HEV-4 transmissions to humans in Europe (Doceul et al., 2016). Coronaviruses infect humans, causing respiratory diseases, and animals. SARS-CoV and MERS-CoV are the sole zoonotic coronaviruses with high pathogenicity in humans that for the last 14 years have caused severe respiratory illness with thousands of cases and deaths. SARS-CoV and MERS-CoV had their evolutionary origin in bats, their natural reservoir, while mammal intermediate hosts, as palm civets for SARS-CoV and camel dromedary for MERS-CoV, are the source of the zoonotic transmission (Lim et al., 2016). Due to the lack of an efficient *in vitro* cultivation system, information on the virus replication, host specificity and evolutionary history are still limited. Surveillance is based on traditional approach (Real-Time RT-PCR or PCR) followed by sequence of short genome fragments conserved among different viral strains. However, this information is not reliable for a robust virus classification and not applicable for the detection of novel viral variants or new viruses (Belak et al., 2009). The Next Generation Sequencing (NGS) technique offers the advantage to generate huge amount of sequence information in one experiment reducing time and cost. In addition, hundreds of sequences from different pathogens or novel variants of known pathogens can be obtained from the same biological sample. For the last years, the NGS applications have grown exponentially, with growing expertise, lower costs and faster results; full genomes can be sequenced for less than € 150 in less than two days. For all of these reasons, full genome sequencing has become an affordable and promising application for research and for public health management (e.g. diagnoses, tracing of outbreaks).

10.1 Hepatitis E

Hepatitis E virus is a major health problem worldwide causing an acute hepatitis that can evolve into chronic forms in immunocompromised patients. Four genotypes infect mammalians. HEV-1 and HEV-2 genotypes infect humans only and cause large epidemic outbreaks in low-income countries. Cases are associated to consumption of drinking water contaminated by human feces, caused by improper release or decontamination of sewages. In industrialized countries, HEV-1 and HEV -2 are reported in travelers returning from endemic areas.

HEV-3 and HEV-4 genotypes are zoonotic and infect both humans and animals. In Europe, more than 20,000 human cases, associated to sporadic cases and small outbreaks of both HEV-3 and HEV-4, have been reported for the last 10 years with an increasing trend. The foodborne transmission in Europe is considered the main via of transmission linked to the consumption of raw or undercooked meat or liver (raw sausages with liver) from pigs and wild boars. HEV-3 is the most common genotype circulating in Europe and has been detected in humans, pigs and wild boar. HEV-4 is prevalent in Asia and has been detected rarely in Europe.

The zoonotic HEV viral strains are characterized by a high genetic heterogeneity and variability. Besides this, the phylogenetic analysis on human and animal strains has showed high evolutionary correlation confirming pigs and wild boars as source of zoonotic transmission. Furthermore, pig and wild boar strains have also showed high correlation hypothesizing the inter-species transmission between livestock and free-range animals or *vice versa*.

10.1.1 Hepatitis E in pigs

The HEV-3 is widespread in pig population worldwide, the seroprevalence rises up to 100% in pig farms all over the world (Pavio et al., 2017). In Italy, the presence of the zoonotic HEV-3 genotype in swine has been frequently reported (Di Bartolo et al., 2008; Masia et al., 2009; Di Martino et al., 2010; Martelli et al., 2010; Monne et al., 2015) and different variants (named subtype) of the genotype HEV-3 have been detected. The most frequently detected subtypes in pigs, as well as humans, are HEV-3f and -3e. Foodborne transmission in sporadic cases and small outbreaks has been confirmed by detection of the same viral strains (100% nucleotide identity) in patients and leftover that has been consumed raw or undercooked. Nevertheless, the biological meaning of subtypes is still unknown; it may also be involved in the evolution of the diseases and in the cross-species transmission.

For the reasons reported above, hepatitis E is now considered an emerging disease also in Europe, and the molecular surveillance of the virus in the animal reservoirs has significantly increased. The surveillance studies are based on the detection of the viral genome by Real Time RT-PCR or nested

RT-PCR, followed by sequencing of short conserved genomic regions such as Methyltransferase (ORF1), RNA-dependent RNA polymerase (ORF1) and the capsid region (ORF2). However, information achieved by sequence comparisons of these regions is limited and may lead to not statistically supported phylogenetic analysis needed for virus classification in subtypes. Comparison of the complete genome sequences is the most appropriate for HEV subtype determination, enabling a more comprehensive classification of the virus.

In the study described in Chapter 3, we identified two swine HEV-3 strains by sequencing of short genomic regions. The analyses highlighted the low correlation of the detected strains with other strains previously detected in Italy and in Europe. The available sequence information was not suitable for a definitely classification in any of the subtypes defined so far. In the same study, to overcome the limit of classification by short genomic sequences, the two swine HEV-3 strains were sequenced by NGS methods, developed in this study, in order to obtain the complete genomes that would be analyzed by phylogenetic analyses. The NGS method was not entirely satisfactory, only few reads were obtained from both samples and the genomes were completed by traditional sequencing methods. The phylogenetic analysis on the complete Italian strain sequences corroborated our previous findings and, following the HEV classification criteria suggested by Smith et al. (2016), a novel HEV-3 subtype, named HEV-3l, was proposed. One pig strain, detected in 2006 in France, clustered within the novel subtype. The cluster was genetically distant from the other HEV strains of both human and animal origin available on NCBI database. Strains belonging to this novel subtype may circulate rarely because have recently emerged in pig population. Changing of subtype circulation in pigs and humans has been observed in some European countries, suggesting changes in the viral circulation that could be justified by animal movements among countries. The heterogeneity of HEV-3 in animal reservoirs may be higher and other subtypes may not be discovered yet. Continuous molecular surveillance is required to monitor circulation of known subtypes and to evaluate occurrence of unknown subtypes that could circulate moderately and emerge in the future.

10.1.2 Hepatitis E in wild boar

In Europe, HEV-3 is frequently reported in wild boars with prevalence varying among countries and ranging between 3.7% (Caruso et al., 2015) and 68.2% (Adlhoch et al., 2009). In Italy, studies reported variable prevalence depending on the geographical origin of the hunted animals and ranging between 1.9% up to 33.7% (Martelli et al., 2008; Serracca et al., 2015). The HEV-3 strains detected in wild boar are almost classified into HEV-3e, -3f, -3c subtypes. The phylogenetic analysis on wild boar strains has showed high evolutionary correlation with swine and human strains, confirming their zoonotic potential (Martelli et al., 2008; Martinelli et al., 2015; Mazzei et

al., 2015; Montagnaro et al., 2015; Serracca et al., 2015; Di Profio et al., 2016; Aprea et al., 2018). In Italy, only few cases are reported every year and the number of available sequences is low. However, the few available reports showed a strict evolutionary correlation among strains identified in human cases and associated to consumption of liver sausages or wild boar meat and HEV-3 strains detected in wild boar (Garbuglia et al., 2015; Mazzei et al., 2015).

As reported above, the origin of infections in humans is unclear, although the suspected HEV sources are domestic pigs but also wild boars. The importance of wild boars as reservoirs of zoonotic HEV-3 is increasingly recognized. As a matter of facts, wild animals are frequently involved in the epidemiology of several zoonoses and, also for HEV, wild animals may serve as reservoirs for transmission to domestic pigs and humans. For this reason, investigating the occurrence of HEV-3 in wild boar, studying the circulation and the origin of the viral strain may help to elucidate pattern of transmission, source attribution and host-jumping.

In Chapters 4 and 6 we reported two surveillance studies on wild boar populations, conducted in Southern and Central Italy (Lazio). Both studies confirmed the circulation of HEV in wild boar revealing a mean prevalence of 14% in South Italy and 52.2% in Central Italy, significantly higher than seroprevalence reported in North Italy (1.5%) (Di Profio et al., 2016). The observed difference may be dependent on the age of the examined animals or to geographical features of the investigated areas (Schielke et al., 2009). The higher prevalence in the Lazio region may be explained by the high density of hunted wild boar, that can be calculated by the number of hunted animals per year; a significantly higher number in the Lazio region, although it is not exactly known. As a matter of facts, the infection is transmitted by oral-fecal route and depends on number of animal sharing the same habitat.

The detected strains were investigated by sequencing and phylogenetic analysis performed on short genomic regions that gave an indication on the subtypes further supported by analyses on the complete genomes obtained by NGS. Results showed circulation of several HEV-3 strains, some of which have been already detected in wild boar in Europe and in Italy such as HEV-3f and HEV-3c. Both subtypes have been detected in human cases and for the last years, a fluctuation of HEV-3 subtypes has been observed in some European countries where the molecular surveillance on human cases is in place (Ijaz et al., 2014; Aspinall et al., 2017). An increase of cases associated to HEV-3c has been observed, although infection associated to the HEV-3efg clade is still the most prevalent. The HEV-3c has also been detected in pigs but in Italy it is more frequent in wild boar, posing the question on the role of this reservoir.

Distribution of HEV-3 subtypes among the investigated areas showed that there was not association with the geographical origin of the investigated animals. The same subtypes were detected in both

areas, such as HEV-3c, while some subtypes were exclusively detected in one of the investigated area (e.g. HEV-3f detected only in animals hunted in the Lazio region).

Our results have highlighted circulation of some uncommon subtypes that have been rarely detected in Europe or never detected before in Italy. This is the case of one wild boar strain WB110VT17, detected in one animal hunted in the Lazio region and classified in a novel provisional subtype recently described in two pigs from Germany and Sweden, and one case of chronic hepatitis E infection in Germany. There are not proof that subtypes and severity of the disease, probably linked to host genetic, are associated. However, studies are needed to ascertain if an association between rarely detected subtypes and occurrence of chronic disease would exist. The novel subtype identified in WB110VT17 is described for the first time in wild boar, confirming the transmission of the same virus among pigs, wild boar and humans.

In conclusion, HEV complete genome sequences from human and other animal reservoirs are required to give new insights on HEV molecular epidemiology and the HEV evolutionary history.

10.2 Coronavirus

Bats are the natural reservoir of several viruses thanks to their biological and ecological features. The capability of flight, affecting the evolution of the immune system and their metabolism, allows them to have a wide distribution worldwide (Calisher et al., 2006; O'Shea et al., 2014; Brook & Dobson, 2015). Bats, infected asymptomatically, are the natural reservoir of Coronaviruses. Indeed, phylogenetic analysis on human and bat CoV strains showed their evolutionary correlation identifying their origin in bats. The monitoring studies on bats CoV have increased worldwide after the SARS and MERS outbreaks since the last decade. Among Betacoronavirus strains, bat SARS-CoV like strains, related to human and palm civet strains, have been described in Rhinolophus bat species, considered the SARS-CoV main reservoir to date. Bat MERS-CoV like strains were detected in African and Chinese bats (Lau et al., 2005; Annan et al., 2013; Ithete et al., 2013; Corman et al., 2014). In three recent studies related strains of the HuCoV-229E have been detected in Hipposideros bats and of HuCoV-NL63 in American tricolored bat (*Perimyotis subflavus*) and Kenyan *Triaenops afer* species suggesting bats as potential reservoir host also for Alphacoronavirus human strains (Pfefferle et al., 2009; Huynh et al., 2012; Corman et al., 2015). In Chapters 7 and 8, we reported the detection and complete genome characterization of Alpha-CoV and Beta-CoV genera strains from two Italian bat species: *Pipistrellus kuhlii* and *Hypsugo savii*. In an ongoing monitoring study, we obtained tissues from one carcasses of *Hypsugo savii* (BatCoV-Ita1), two carcasses (BatCoV-Ita2, BatCoV-Ita3) and two intestinal contents (BatCoV-Ita4, BatCoV-Ita5) from *Pipistrellus kuhlii*, that were positive for the presence of CoV. In Italy, these species colonise

and forage in different environment. In particular, *Pipistrellus kuhlii* bats colonise agricultural and urban areas while *Hypsugo savii bats* feed at lights in rural areas, towns and cities.

The random NGS sequencing approach was used to obtain the CoV complete genomes from bat biological samples in order to classify and characterize completely the viruses. The ICTV classification criteria based on the 90% amino acid sequence identity threshold values in the conserved concatenated domains of orf1ab polyprotein was applied. From the tissues of *Pipistrellus kuhlii* and *Hypsugo savii*, two almost identical Betacoronavirus strains were retrieved and classified into the MERS-CoV species. The phylogenetic analysis showed the evolutionary relationship to MERS-CoV and MERS-CoV like strains sharing a common ancestor. Interestingly the two MERS-CoV like strains detected were found in two species that lives in different environments. Despite the different habitats, two strains, showing high evolutionary correlation (99% nt. id.), have been adapted in two different bat species belonging to the *Vespertilionidae* family. In the intestinal contents and one carcass of three *Pipistrellus kuhlii* bats, three Alphacoronavirus strains, divergent from all other Alpha-CoV species, were retrieved. Following the ICTV demarcation criteria two novel species were established. The phylogenetic analysis showed that one Alpha-CoV species have been detected only in *Pipistrellus kuhlii* species in Europe, suggesting that this species may have been adapted to this host only. The other Alpha-CoV species have been related to an Italian and to a French strain collected in 2014 and detected in Pipistrellus Pipistrellus. The detection of one CoV species in two bat species (*Pipistrellus kuhlii* and *Pipistrellus Pipistrellus*) indicates the ability of this species to infect different hosts of the same *Pipistrellus* genera.

Since the last decade the surveillance studies on CoVs in bats have increased, however, few studies only have reported the complete genome sequences of the strains. Up to date, the number of CoV species, infecting bat that are reported with complete genome description are 9. Six from Alpha-CoV genera: *Miniopterus bat coronavirus 1*, *Bat coronavirus CDPHE15*, *Miniopterus bat coronavirus HKU8*, *Rhinolophus bat coronavirus HKU2*, *Bat coronavirus HKU10*, *Scotophilus bat coronavirus 512.* Three from Beta-CoV genera: *Bat coronavirus HKU4, Bat coronavirus HKU5, Bat coronavirus HKU9 (*https://talk.ictvonline.org/taxonomy/). The number of CoV species increases if we consider the classification based on short genomic sequences. Moreover, the heterogeneity of CoVs in bats may be higher if we consider that a minor percentage only (20%) of the known bat species has been monitored for the presence of CoVs. Although we do not have a clear understanding of the circulation of CoV species between bat species, additional monitoring studies on CoV in bats will clarify their evolution and important mechanisms in order to cross the species barrier.

10.3 NGS methods

Since the last decade, Next Generation Sequencing (NGS), known as high-throughput sequencing, has increased the speed and accuracy of produced data, compared to the classical sequencing methods (Shendure & Ji, 2008). This novel technique has allowed the discovering of novel viruses, tracking of outbreaks and pandemic events. The detection of the sequence of interest in the whole reads obtained from a sample is a major problem in the NGS method and it also depends on the quality and preparation of the samples (Barzon et al., 2011). Because of this, NGS has been applied rarely in HEV sequencing. The main obstacle to HEV complete genome amplification is the low viral load in feces, food or environmental samples. In the absence of an efficient virus cell culture system, only molecular approach can be applied to enrich the virus nucleic acids.

To enrich the viral nucleic acids in the sample for NGS, three main steps are important: sample preparation, library preparation and data analysis. In our study, the three steps have been implemented. The virus nucleic acids enrichment was obtained during the sample preparation by filtration to remove bacteria and by DNase and RNase treatment to digest away cellular nucleic acids and enrich the capsid-protected viral nucleic acids (Delwart, 2007). The library preparation of the nucleic acids from biological samples was obtained using random retro-transcription and amplification techniques, Sequence Independent Single Primer Amplification (SISPA) (Djikeng et al., 2008) (used in Chapter 3) usually used for metagenomics, and using multiplexed PCR targeted sequences (Chapter 9). Finally, during data analysis the bioinformatic pipeline was developed and uploaded onto the online Galaxy tool to perform data quality check of the obtained reads (sequences), to remove host and bacteria nucleic acids avoiding false positive alignment and to assemble the complete genome by mapping the best quality reads to a reference genome. In the first study, the SISPAS was firstly applied to HEV positive fecal samples to obtain HEV complete genomes as described in Chapter 3. Despite the high number of sequences obtained from the NGS run, only 0.1% were specific HEV and not enough to obtain the complete genome sequence. This could be due to the low HEV viral load in the sample 2.8×10^5 (SWHEV75BO2012) and 7.1×10^5 (HEV/13RS985-5) HEV genome equivalents (GE), respectively and due to the presence of nucleic acids of other viruses (10% of total reads), of bacteria and host (90%), that were still present despite the sample treatment before RNA extraction. The same NGS protocol (SISPA) was applied to HEV positive liver sample as described in Chapter 4. One strain (WB/HEV/NA21ITA15) was completely sequenced and one (WB/HEV/NA17ITA15) lacked the first 20 nt at 5' terminus region. The NGS run resulted in a limited number of reads, of which 70% were from the host. However, the high HEV-3 load in the liver ($3.8x10^6$ and $1.7x10^8$ GE) enabled to recover the full genome sequence using the reads obtained. As expected, results of NGS metagenomic approach to obtain full genome

depended mainly on the load of the target nucleic acids in the sample that was only moderately enriched by physical and enzymatic treatments. Furthermore, the HEV-3 complete genomes were obtained with a low mean coverage rate (<25x).

While the NGS technology is less used for HEV whole genome sequencing and metagenomics analysis, more studies have reported the application of this technology on CoV positive samples. The NGS technology was used to obtain partial or complete CoV genome sequences for metagenomics or phylogenetic and characterization analysis. Alpha-CoV and Beta-CoV genera strains were amplified from bat (Yang et al., 2013; Mendenhall et al., 2017; Geldenhuys et al., 2018), human (Cotten et al., 2013), camel (Briese et al., 2014; Yusof et al., 2017), swine, avian and other animal species. In our study, the random NGS protocol was applied to carcasses and fecal samples from Italian bats. The protocol resulted successful to obtain the complete genome. Since the RNA only was available in our laboratory, no treatments were applied on the biological samples. In addition, despite the high presence of non-viral sequences in the sample, a great number on CoV sequences were retrieved (<30% of the total reads) and sufficient to obtain the complete genome with high coverage rate (<2000x). Since no quantitative Real Time RT-PCR was available to quantify the CoV RNA, we could hypothesize that the viral load was high enough or at least higher than HEV in fecal sample, to obtain thousands of CoV sequences.

However, to obtain HEV and CoVs complete genome, almost 1 million reads per sample are needed to obtain sufficient number of specific reads to complete the entire genome. In this study, the Ion Personal Machine (PGM) sequencer was used. The libraries were pool and loaded on Ion 318 v2 Chip, a physical support where 4-5.5 million reads per run can be sequenced. Considering the minimum number of sequences to obtain HEV and CoV genomes, 2 HEV or 4 CoV samples can be loaded on a Ion 318 v2 Chip. The number of samples that can be analyzed limits the use of this approach for diagnoses or surveillance. To increase the efficacy of NGS, the accuracy of sequence recovered, increasing the sequence coverage, and to reduce the cost of experiments, a multiplexed PCR enrichment method was developed as described in Chapter 9. The two primer panels designed for HEV and MERS were successful, allowing the amplification of the complete genome of the tested samples. Using this method, a significant increase of specific viral reads was observed and compared to previous results. An increasing of specific sequences leads to a high coverage rate using also a number of total reads lower than the random approach. In fact, to obtain a HEV or CoV complete genome with a coverage of >1000x, more than 25 samples can be loaded on a single chip, reducing the cost to €150 per sample. The reduced time needed for library preparation and the effective cost make this novel approach applicable to clinical diagnostic, surveillance studies or during an outbreak.

10.4 Concluding remarks

The work described in this PhD thesis focused on the detection, for HEV, and characterization of both zoonotic viruses, HEV and Coronaviruses, in their respective animal reservoir using classical sequencing technique and Next Generation Sequencing. The detection based on classical technique cannot allow the detection of novel viral variants. For this reason, novel NGS methods were applied. The NGS technique resulted a powerful and promising method to detect and sequence viral complete genomes and different variants of the zoonotic HEV-3 and MERS-CoV strains. The first developed random approach was not totally efficient, however, it allowed to obtain partial genome sequences resulting useful for further studies and metagenomic analysis on biological samples. The target amplification method was totally satisfactory, allowing to obtain complete genomes with high coverage rate and to detect several viral variants within a single NGS run in less than two days. The molecular characterization of the obtained complete genomes by phylogenetic analysis resulted in a better classification of the virus, with a strong statistical support (bootstrap method). In addition, an accurate classification allows a more detailed comprehension of the strains circulating in animal population, detecting novel variants, to which human may be exposed. The detected viruses may be correlated to human strains under the evolutionary point of view, showing a zoonotic potential, while other strains can by highly divergent to the currently known ones. The results presented in this thesis provide new knowledge, filling some of the gaps in HEV and CoV evolution. In addition, the NGS methods developed during the PhD period and described in this thesis, will be useful also in future studies for the monitoring of these viruses with an increased throughput, and will be relevant in the case of outbreaks, to detect and characterize novel virus strains.

10.5 References

Adlhoch, C., Wolf, A., Meisel, H., Kaiser, M., Ellerbrok, H. and Pauli, G., 2009: High HEV presence in four different wild boar populations in East and West Germany. *Veterinary Microbiology,* **139,** 270-278.

Annan, A., Baldwin, H.J., Corman, V.M., Klose, S.M., Owusu, M., Nkrumah, E.E., et al., 2013: Human Betacoronavirus 2c EMC/2012-related Viruses in Bats, Ghana and Europe. *Emerging Infectious Diseases,* **19,** 456-459.

Aprea, G., Amoroso, M.G., Di Bartolo, I., D'Alessio, N., Di Sabatino, D., Boni, A., et al., 2018: Molecular detection and phylogenetic analysis of hepatitis E virus strains circulating in wild boars in south-central Italy. *Transboundary and emerging diseases,* **65,** e25-e31.

Aspinall, E.J., Couturier, E., Faber, M., Said, B., Ijaz, S., Tavoschi, L., et al., 2017: Hepatitis E virus infection in Europe: surveillance and descriptive epidemiology of confirmed cases, 2005 to 2015. *Eurosurveillance,* **22,** 13-22.

Barzon, L., Lavezzo, E., Militello, V., Toppo, S. and Palu, G., 2011: Applications of Next-Generation Sequencing Technologies to Diagnostic Virology. *International journal of molecular sciences,* **12,** 7861-7884.

Belak, S., Thoren, P., LeBlanc, N. and Viljoen, G., 2009: Advances in viral disease diagnostic and molecular epidemiological technologies. *Expert Review of Molecular Diagnostics,* **9,** 367-381.

Briese, T., Mishra, N., Jain, K., Zalmout, I.S., Jabado, O.J., Karesh, W.B., et al., 2014: Middle East Respiratory Syndrome Coronavirus Quasispecies That Include Homologues of Human Isolates Revealed through Whole-Genome Analysis and Virus Cultured from Dromedary Camels in Saudi Arabia. *Mbio,* **5**.

Brook, C.E. and Dobson, A.P., 2015: Bats as 'special' reservoirs for emerging zoonotic pathogens. *Trends in Microbiology,* **23,** 172-180.

Calisher, C.H., Childs, J.E., Field, H.E., Holmes, K.V. and Schountz, T., 2006: Bats: Important reservoir hosts of emerging viruses. *Clinical Microbiology Reviews,* **19,** 531-+.

Caruso, C., Modesto, P., Bertolini, S., Peletto, S., Acutis, P.L., Dondo, A., et al., 2015: Serological and virological survey of hepatitis E virus in wild boar populations in northwestern Italy: detection of HEV subtypes 3e and 3f. *Archives of Virology,* **160,** 153-160.

Corman, V.M., Baldwin, H.J., Tateno, A.F., Zerbinati, R.M., Annan, A., Owusu, M., et al., 2015: Evidence for an Ancestral Association of Human Coronavirus 229E with Bats. *Journal of Virology,* **89,** 11858-11870.

Corman, V.M., Ithete, N.L., Richards, L.R., Schoeman, M.C., Preiser, W., Drosten, C., et al., 2014: Rooting the Phylogenetic Tree of Middle East Respiratory Syndrome Coronavirus by Characterization of a Conspecific Virus from an African Bat. *Journal of Virology,* **88,** 11297-11303.

Cotten, M., Lam, T.T., Watson, S.J., Palser, A.L., Petrova, V., Grant, P., et al., 2013: Full-Genome Deep Sequencing and Phylogenetic Analysis of Novel Human Betacoronavirus. *Emerging Infectious Diseases,* **19,** 736-742.

Delwart, E.L., 2007: Viral metagenomics. *Reviews in Medical Virology,* **17,** 115-131.

Di Bartolo, I., Martelli, F., Inglese, N., Pourshaban, M., Caprioli, A., Ostanello, F., et al., 2008: Widespread diffusion of genotype 3 hepatitis E virus among farming swine in Northern Italy. *Veterinary Microbiology,* **132,** 47-55.

Di Martino, B., Di Profio, F., Martella, V., Di Felice, E., Di Francesco, C.E., Ceci, C., et al., 2010: Detection of hepatitis E virus in slaughtered pigs in Italy. *Archives of Virology,* **155,** 103-106.

Di Profio, F., Melegari, I., Sarchese, V., Robetto, S., Marruchella, G., Bona, M.C., et al., 2016: Detection and genetic characterization of hepatitis E virus (HEV) genotype 3 subtype c in wild boars in Italy. *Archives of Virology,* **161,** 2829-2834.

Djikeng, A., Halpin, R., Kuzmickas, R., DePasse, J., Feldblyum, J., Sengamalay, N., et al., 2008: Viral genome sequencing by random priming methods. *Bmc Genomics,* **9**.

Doceul, V., Bagdassarian, E., Demange, A. and Pavio, N., 2016: Zoonotic Hepatitis E Virus: Classification, Animal Reservoirs and Transmission Routes. *Viruses-Basel,* **8**.

Dong, J., Olano, J.P., McBride, J.W. and Walker, D.H., 2008: Emerging pathogens: Challenges and successes of molecular diagnostics. *The Journal of Molecular Diagnostics,* **10,** 185-197.

Garbuglia, A.R., Alessandrini, A.I., Pavio, N., Tesse, S., Grignolo, S., Viscoli, C., et al., 2015: Male patient with acute hepatitis E in Genoa, Italy: figatelli (pork liver sausage) as probable source of the infection. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases,* **21,** e4-6.

Geldenhuys, M., Mortlock, M., Weyer, J., Bezuidt, O., Seamark, E.C.J., Kearney, T., et al., 2018: A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in Neoromicia bats within South Africa. *PloS one,* **13,** e0194527.

Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J.C., et al., 2012: Evidence supporting a zoonotic origin of human coronavirus strain NL63. *Journal of Virology,* **86,** 12816-12825.

Ijaz, S., Said, B., Boxall, E., Smit, E., Morgan, D. and Tedder, R.S., 2014: Indigenous hepatitis E in England and wales from 2003 to 2012: evidence of an emerging novel phylotype of viruses. *The Journal of Infectious Diseases,* **209,** 1212-1218.

Ithete, N.L., Stoffberg, S., Corman, V.M., Cottontail, V.M., Richards, L.R., Schoeman, M.C., et al., 2013: Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. *Emerging Infectious Diseases,* **19,** 1697-1699.

Karesh, W.B., Dobson, A., Lloyd-Smith, J.O., Lubroth, J., Dixon, M.A., Bennett, M., et al., 2012: Zoonoses 1 Ecology of zoonoses: natural and unnatural histories. *Lancet,* **380,** 1936-1945.

Lau, S.K.P., Woo, P.C.Y., Li, K.S.M., Huang, Y., Tsoi, H.W., Wong, B.H.L., et al., 2005: Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proceedings of the National Academy of Sciences USA,* **102,** 14040-14045.

Lim, Y.X., Ng, Y.L., Tam, J.P. and Liu, D.X., 2016: Human Coronaviruses: A Review of Virus-Host Interactions. *Diseases,* **4**.

Martelli, F., Caprioli, A., Zengarini, M., Marata, A., Fiegna, C., Di Bartolo, I., et al., 2008: Detection of hepatitis E virus (HEV) in a demographic managed wild boar (Sus scrofa scrofa) population in Italy. *Veterinary Microbiology,* **126,** 74-81.

Martelli, F., Toma, S., Di Bartolo, I., Caprioli, A., Ruggeri, F.M., Lelli, D., et al., 2010: Detection of Hepatitis E Virus (HEV) in Italian pigs displaying different pathological lesions. *Research in veterinary science,* **88,** 492-496.

Martinelli, N., Pavoni, E., Filogari, D., Ferrari, N., Chiari, M., Canelli, E., et al., 2015: Hepatitis E virus in wild boar in the central northern part of Italy. *Transboundary and emerging diseases,* **62,** 217-222.

Masia, G., Orru, G., Liciardi, M., Desogus, G., Coppola, R.C., Murru, V., et al., 2009: Evidence of hepatitis E virus (HEV) infection in human and pigs in Sardinia, Italy. *Journal of preventive medicine and hygiene,* **50,** 227-231.

Mazzei, M., Nardini, R., Verin, R., Forzan, M., Poli, A. and Tolari, F., 2015: Serologic and molecular survey for hepatitis E virus in wild boar (Sus scrofa) in Central Italy. *New microbes and new infections,* **7,** 41-47.

Mendenhall, I.H., Borthwick, S., Neves, E.S., Low, D., Linster, M., Liang, B., et al., 2017: Identification of a Lineage D Betacoronavirus in Cave Nectar Bats (Eonycteris spelaea) in Singapore and an Overview of Lineage D Reservoir Ecology in SE Asian Bats. *Transboundary and emerging diseases,* **64,** 1790-1800.

Monne, I., Ceglie, L., Di Martino, G., Natale, A., Zamprogna, S., Morreale, A., et al., 2015: Hepatitis E virus genotype 4 in a pig farm, Italy, 2013. *Epidemiology and infection,* **143,** 529-533.

Montagnaro, S., De Martinis, C., Sasso, S., Ciarcia, R., Damiano, S., Auletta, L., et al., 2015: Viral and Antibody Prevalence of Hepatitis E in European Wild Boars (Sus scrofa) and Hunters at Zoonotic Risk in the Latium Region. *Journal of Comparative Pathology,* **153,** 1-8.

O'Shea, T.J., Cryan, P.M., Cunningham, A.A., Fooks, A.R., Hayman, D.T.S., Luis, A.D., et al., 2014: Bat Flight and Zoonotic Viruses. *Emerging Infectious Diseases,* **20,** 741-745.

Pavio, N., Doceul, V., Bagdassarian, E. and Johne, R., 2017: Recent knowledge on hepatitis E virus in Suidae reservoirs and transmission routes to human. *Veterinary research,* **48**.

Pfefferle, S., Oppong, S., Drexler, J.F., Gloza-Rausch, F., Ipsen, A., Seebens, A., et al., 2009: Distant Relatives of Severe Acute Respiratory Syndrome Coronavirus and Close Relatives of Human Coronavirus 229E in Bats, Ghana. *Emerging Infectious Diseases,* **15,** 1377-1384.

Schielke, A., Sachs, K., Lierz, M., Appel, B., Jansen, A. and Johne, R., 2009: Detection of hepatitis E virus in wild boars of rural and urban regions in Germany and whole genome characterization of an endemic strain. *Virology Journal,* **6**.

Serracca, L., Battistini, R., Rossini, I., Mignone, W., Peletto, S., Boin, C., et al., 2015: Molecular Investigation on the Presence of Hepatitis E Virus (HEV) in Wild Game in North-Western Italy. *Food and Environmental Virology,* **7,** 206-212.

Shendure, J. and Ji, H.L., 2008: Next-generation DNA sequencing. *Nature Biotechnology,* **26,** 1135-1145.

Smith, D.B., Simmonds, P., Izopet, J., Oliveira, E.F., Ulrich, R.G., Johne, R., et al., 2016: Proposed reference sequences for hepatitis E virus subtypes. *The Journal of General Virology,* **97,** 537-542.

Yang, L., Wu, Z., Ren, X., Yang, F., He, G., Zhang, J., et al., 2013: Novel SARS-like betacoronaviruses in bats, China, 2011. *Emerging Infectious Diseases,* **19,** 989-991.

Yusof, M.F., Queen, K., Eltahir, Y.M., Paden, C.R., Al Hammadi, Z., Tao, Y., et al., 2017: Diversity of Middle East respiratory syndrome coronaviruses in 109 dromedary camels based on full-genome sequencing, Abu Dhabi, United Arab Emirates. *Emerging microbes & infections,* **6,** e101.