UNIVERSITA' DEGLI STUDI DI MILANO
Department of Health, Animal Science and Food Safety

PhD Course in Biotechnologies applied to veterinary and animal sciences
Class XXVIII

SPATIAL AND TEMPORAL DYNAMICS OF PESTIVIRUS, CRIMEAN-CONGO HEMORRHAGIC FEVER VIRUS AND HEPATITIS E VIRUS: A PHYLOGENETIC APPROACH

Tutor Prof. Valerio Bronzo
Co-tutor Prof. Gianguglielmo Zehender

PhD Coordinator: Prof. Fulvio Gandolfi

PhD Candidate Camilla Luzzago
R09997

Academic Year 2014-2015

A Paolo, Attilio, Ernesto, Augusto

CONTENTS

ABSTRACT

ABSTRACT

The general aim of the PhD project was to apply phylogenetic analysis to viral sequences obtained in different geographic areas at different times, in order to reconstruct the most probable places of origin and pathways of dispersion of infections. Viral population growth and evolution leave a measurable imprint on the genome of viruses over the course of years, months or even days and occur simultaneously with geographic dispersal (Holmes, 2008; Pybus & Rambaut, 2009). This interaction characterizes a spatial phylodynamic process that can be recovered from genomic data using phylogeographic analyses (Faria et al., 2011). The research activity has been focused on Pestivirus genus, that includes pathogens of livestock (Bovine viral diarrhea virus - BDV) and wildlife (Border disease virus - BDV), and on zoonotic emerging diseases, involving in their epidemiological cycle both livestock and wildlife (Crimean-Congo hemorrhagic fever - CCHF, Hepatitis E virus- HEV).

Concerning BDV, since 2001 several outbreaks of disease have been reported in Pyrenean chamois in Spain, France and Andorra. These outbreaks have decimated several Pyrenean chamois populations, with mortalities ranging from 40% to 85%. The infection has become endemic in the Central and Eastern Pyrenees. The aim of this study was to clarify the origin and dispersion of the Pyrenean chamois BDV genetic variant by reconstructing the spatial and temporal dynamics of BDV 5' UTR sequences of Pyrenean chamois, 10 novel sequences and 41 retrieved from public databases and Sheep BDV sequences (n=44) from Spain and France were also retrieved. The phylogenetic analysis was performed using a Bayesian Markov chain Monte Carlo (MCMC) method implemented in the BEAST v.1.74 package. The chamois clade

originated from sheep BDV genotype 4, generating a founder effect due to intra-species spread and spatial dispersion. The time of the most recent common ancestor estimates for the chamois clade dated back to a time span between 1974 and 1996, with a mean estimation falling in 1988. The pathway of dispersion of isolates suggests a complex exchange between neighboring Pyrenean chamois populations, still going on such as Western direction.

Genetic typing of bovine viral diarrhea virus (BVDV) has distinguished BVDV-1 and BVDV-2 species and an emerging putative third species (HoBi-like virus), recently detected in southern Italy, signaling the occurrence of natural infection in Europe. Recognizing the need to update the data on BVDV genetic variability in Italy for mounting local and European alerts, a wide collection of 5' UTR sequences (n = 371) was selected to identify the frequency of genotypes and subtypes at the herd level. BVDV-1 had the highest frequency, followed by sporadic BVDV-2. No novel HoBi-like viruses were identified. Four distribution patterns of BVDV-1 subtypes were observed: highly prevalent subtypes with a wide temporal-spatial distribution (1b and 1e), low prevalent subtypes with a widespread geographic distribution (1a, 1d, 1g, 1h, and 1k) or a restricted geographic distribution (1f), and sporadic subtypes detected only in single herds (1c, 1j, and 1l). BVDV-1c, k, and l are reported for the first time in Italy. Italy is one of the countries with the highest genetic diversity of BVDV worldwide. Northern Italy ranked first for BVDV introduction, prevalence, and dispersion. Nevertheless, the presence of sporadic variants in other restricted areas suggests the risk of different routes of BVDV introduction.

CCHF is a zoonosis mainly transmitted by ticks that causes sporadic cases and severe hemorrhagic fever of acute human disease with a mortality rate of 5-60% and it has recently emerged in the Balkans and eastern Mediterranean areas. In order to reconstruct the origin and pathway of the worldwide dispersion of the virus at global and regional (eastern European) level, we investigated the phylogeography of the infection by analysing 121 publicly available CCHFV S gene sequences including two recently characterised Albanian isolates. The spatial and temporal phylogeny was reconstructed using a Bayesian Markov chain Monte Carlo approach.

CCHFV phylogeographic reconstruction suggests that the disease originated about one thousand years ago from a common ancestor probably located in Africa. The virus then spread to Asia in the XV century and entered Europe on at least two occasions: the first in the early 1800s and the second in the early 1900s. The most probable location for the origin of the European clade was Russia, but Turkey played a central role in spreading the virus throughout Europe. Our data suggest that the movement of wild and domestic ungulates from endemic areas probably represent the main cause of virus dissemination in Eastern Europe.

Hepatitis E virus is classified into four genotypes that have different geographical and host distributions. The main cause of sporadic autochthonous type E acute hepatitis in developed countries is genotype 3, which has a worldwide distribution and widely infects pigs. The aim of this study was to make hypotheses concerning the origin and global dispersion routes of this genotype by reconstructing the spatial and temporal dynamics of 208 HEV genotype 3 ORF-2 sequences (retrieved from public databases) isolated in different geographical areas. The evolutionary rates, time of the most

recent common ancestors (tMRCAs), epidemic growth and phylogeography of HEV-3 were co-estimated using a MCMC Bayesian method. On the basis of time-scaled phylogeographical reconstruction, we hypothesise that HEV-3, after originating in the early 1800s in Europe, reached Asia in the first decades of 1900, and then moved to America probably in the 1970s-1980s. Analysis of the skyline plot showed a sharp increase of the number of infections between the 1980s and 2005, suggesting the intervention of new and highly efficient routes of transmission, possibly related to changes in the pig industry.

INTRODUCTION

Scientific background

The understanding of the origin, the study of the adaptation speed and of the selective factors give important information on the epidemiology of pathogens and on possible prevention and control measures. The advanced phylogenetic analysis gives these information through the reconstruction of the evolutionary history of the pathogens. These analytical methods are innovative for human medicine and even more for the veterinary sciences. Among the advanced methods for the molecular evolution analysis, those estimating the evolutionary rates of phylogenetically correlated isolates, play an important role. Sequences sampled at different points in time (heterochronous), among which there are a statistically significant number of genetic differences, allow to study the evolutionary process in a calendar timescale (Drummond et al., 2002). These methods are particularly useful in the study of rapidly evolving viruses which can accumulate a great amount of mutations in a short sampling time, such as the majority of RNA and some DNA viruses (Duffy et al., 2008). Maximum-likelihood or Bayesian methods which allow to estimate the evolutionary rate of dated sequences into a phylogenetic tree have been developed. The possibility to infer the epidemiology history of an infection by phylogeny of a samples of isolates gave rise to a new framework called “phylodynamics” (Grenfell et al., 2004). Several methods have been developed to reconstruct the spatial structure of a phylogeny on the basis of the localities of origin of the isolates used for the analysis (phylogeography) and recently applied to viral zoonosis such as rabies (Talbi et al., 2010) and West Nile Virus (Zehender et al., 2011), highlighting that phylogeographic approach contributes to characterizing and predicting the spatial spread of pathogens.

Emerging and re-emerging virus infections

Emerging infectious diseases, most of which are caused by RNA viruses, have been reported with an increasing importance for public health in recent years (Cleaveland et al., 2001). The typically broader host ranges and much higher mutation rates of RNA viruses facilitate both the initial infection of a new host and subsequence adaptation to that host (Woolhouse et al., 2001, 2005).

The first vertebrate virus described, an aphthovirus of the family *Picornaviridae*, that causes foot and mouth disease, was isolated from a cow in 1897 (Loeffler & Frosch, 1897). Since then, the rate at which virulent viruses have been discovered has been determinate by the ability of available technology to identify viruses, and the ability to collect specimens potentially containing novel viruses. The mean annual rate of pathogenic virus discovery showed an increase during 1960–1969 to 4.9/year, after the introduction of diagnostic innovation and thanks to intensive activities of tropical field research stations, followed by a sudden deceleration in the rate of discovery to only about 2/year, which persisted through 2010 (Rosenberg, 2015). The availability of increasingly powerful methods for genomic characterization, such as next-generation sequencing (NGS), also known as high-throughput sequencing, will likely contribute to detect and characterize pathogens without prior knowledge of their existence, but mainly viruses that have mutualistic or symbiotic relationships with their hosts (Roossinck, 2011).

Approximately 60% of infectious diseases in humans, and most of recent pandemic threats, have had a zoonotic origin (Woolhouse et al., 2005, Cutler et al., 2010). RNA viruses are predominant (Woolhouse & Gaunt, 2007), mostly due to the rate of error

during RNA replication, equivalent to nucleotide substitution rates of 10^{-3} to 10^{-5} substitutions per site per year (subs/site/year), compared to a 10^{-7} to 10^{-9} subs/site/year of dsDNA (Duffy et al., 2008). The range corresponds to the fidelity of the polymerases used in replication: RNA viruses which utilize RNA-dependent RNA polymerases mutate faster than retroviruses (with RNA-dependent DNA polymerases) or reverse transcriptases which mutate faster than DNA viruses (with DNA polymerases).

For the emerging and reemerging pathogen species, the main drivers have been recently systematic reviewed by Woolhouse & Gowtage-Sequeria (2005) (table 1). The overall conclusions drawn by the authors highlighted that although some drivers, such as changes in land use or agricultural practices, are more important for zoonotic diseases than for non zoonotic ones, pathogens are exploiting almost any change in human ecology that provides new opportunities for transmission, either between humans or to humans from a nonhuman source.

Actually, emerging infectious diseases are often the result of a pathogen jumping from natural host into a new host species. The process of host jumping is of particular importance for RNA viruses, their rapid evolutionary dynamics provide them with the capacity to quickly generate the genetic variation required to evade both innate and adaptive host immunity (Holmes, 2008). The risk of a host shift will depend on the likelihood that the necessary set of mutations can accumulate in the newly infected host (Longdon et al., 2014). Host jumping or host shift can cause spillover events that result in dead end infections in the new host, such as for example the Hendra and Nipah viruses and avian influenza viruses in human (Chua et al., 2000; Webby &

Webster, 2001), or a host shift with successful transmission in the new host, such as SARS coronavirus (Li et al., 2005).

Table 1 Main categories of drivers associated with emergence and reemergence of human pathogens (Woolhouse & Gowtage-Sequeria, 2005)

Rank*	Driver
1	Changes in land use or agricultural practices
2	Changes in human demographics and society
3	Poor population health (e.g., HIV, malnutrition)
4	Hospitals and medical procedures
5	Pathogen evolution (e.g., antimicrobial drug resistance, increased virulence)
6	Contamination of food sources or water supplies
7	International travel
8	Failure of public health programs
9	International trade
10	Climate change

*Ranked by the number of pathogen species associated with them (most to least).

RNA viruses and pathogens that already have a broad host range being particularly prone to jumping between distantly related species (Woolhouse et al., 2005; Woolhouse & Gowtage-Sequeria, 2005).

Concerning the variation in susceptibility of a new host species to a new pathogen, closely related species to the natural host of pathogen are the most susceptible, defined as ‘‘phylogenetic distance effect’’ (Longdon et al., 2014). Closely related species may also have similar levels of susceptibility, regardless of their distance from the pathogen’s natural host, defined as “phylogenetic clade effect” (Longdon et al., 2014), such as the recurrently shift between distantly related taxa in influenza virus among birds, pigs, and humans (Webby & Webster, 2001)

One of the most recent emergence of RNA virus is represented by Middle East respiratory syndrome (MERS), a highly lethal respiratory disease in human caused by a novel single-stranded, positive-sense RNA betacoronavirus (MERS-CoV), the first case identified in June 2012). As recently reviewed by Zumla et al., (2015), dromedary camels, hosts for MERS-CoV, are implicated in direct or indirect transmission to human beings, although the exact mode of transmission is unknown. As of May 31, 2015, 1180 laboratory-confirmed cases of MERS and 483 deaths have been reported to WHO. Although most cases of MERS have occurred in Middle East (namely Saudi Arabia, the United Arab Emirates), sporadic cases have been reported in Europe, the USA, and Asia in people who travelled from the endemic areas. Little human-to-human transmission reported in the community. MERS-CoV continues to be an endemic, low-level public health threat.

Evolutionary analysis of the dynamics of viral infectious diseases

The field of viral evolutionary analysis has greatly benefited from three developments: the increasing availability and quality of viral genome sequences, the growth in

computer processing power, and the development of sophisticated statistical methods (Pybus & Rambaut, 2009).

The study of the origins, emergence, and spread of viral infections in populations is one of the most active and productive areas of research in modern evolutionary biology (Holmes, 2008). Actually, RNA virus molecular evolution is usually occurring at approximately the same rate as the ecological processes that shape their diversity, viral phylogenies will often contain information about temporal as well as spatial dynamics (Holmes, 2004), therefore genetic variation generated can be used to infer the patterns, processes, and dynamics of viral evolution.

Phylodynamic analyses represent the relationship between genetic distance and time, the viral sequences have to be sampled at known time points, this approach estimates the ages of branching events and the time of the common ancestor (Pybus & Rambaut, 2009), assuming a constant rate of variation of virus evolution (strict clock model) or more complex models that assume an evolutionary rates variation through time and among lineages (relaxed clock models).

Phylogeography combined spatial and temporal analyses to reveal location and time of origin of infections, route of transmission and flow of geographic spread (Pybus & Rambaut, 2009). The term phylogeography is related to the principles and processes governing the geographical distributions of genealogical lineages, especially those at the intraspecific level (Avise, 1998) and the general aim of phylogeography might be defined as a means to understand microevolution and speciation in its geographic or spatiotemporal context (Kidd & Ritchie, 2006). Pathways of spatial distribution are fixed in phylogenetic trees as a record of changes between geographic locations along

tree branches and phylogeographic approaches help uncover the imprint that spatial epidemiological processes leave in the genomes of fast evolving viruses (Faria et al., 2011). Several methods have been developed to reconstruct the spatial structure of a phylogeny on the basis of the localities of origin of the isolates used for the analysis and recently it has been develop a Bayesian framework for inference, visualization and hypothesis testing of phylogeographic history in a statistically efficient fashion (Lemey et al., 2009), that demonstrated several advantages of simultaneously inferring spatial and temporal processes from gene sequences.

Phylogeography analysis has been recently applied to rabies both in endemic areas in North Africa and also in re-emergent outbreaks in wildlife in Italy. Rabies is an important zoonotic disease, that every year caused approximately 55,000 human deaths (Knobel et al, 2005), almost all lethal cases occurring in developing countries where rabies virus (RABV; negative-sense RNA virus, family *Rhabdoviridae*) is endemic in the domestic dog (Cleaveland et al., 2006). Despite the importance of dogs as vectors for human rabies, little is known about the spatial and temporal dynamics of rabies in this major reservoir species, or the processes responsible for its maintenance in specific geographic localities (Talbi et al., 2009), therefore the authors investigated how evolutionary dynamics shape the spatial distribution and spread of rabies in dog in North Africa, suggesting that the human mediated dispersal of infected dogs is likely to continue to play a major role in the transmission of RABV in geographical areas where it has been present for many years.

Italy has been classified as free from terrestrial rabies since 1997, the last case of rabies was diagnosed in a fox in the province of Trieste on the border with Slovenia in

December 1995. In late 2008, fox rabies re-emerged in North-eastern Italy, with a total of 287 rabies cases confirmed in animals by February 2011. As with previous rabies epidemics, the recent Italian outbreaks have been linked to the epidemiological situation in the adjacent Balkan region (De Benedictis et al., 2008). Despite the implementation of four oral rabies vaccination (ORV) campaigns the infection spread westwards (De Benedictis et al., 2009; Nouvellet et al., 2013). Fusaro et al. (2013) suggest that the re-emergence of fox rabies into North-eastern Italy was due to two viral genetic groups introduced a few months apart and occupied two distinct geographic areas, Italy-1 were likely introduced from Slovenia or Croatia while group Italy-2 appears to have been introduced from Bosnia and Herzegovina and concluded that the lack of adequate and homogeneous surveillance, particularly at border areas, has clearly increased the risk of introduction of wildlife rabies.

AIMS

AIMS

The general aim was to apply phylogenetic analysis to viral sequences obtained in different geographic areas at different times, in order to reconstruct the most probable places of origin and pathways of dispersion of infections. Viral population growth and evolution leave a measurable imprint on the genome of viruses over the course of years, months or even days and occur simultaneously with geographic dispersal (Holmes, 2008; Pybus & Rambaut, 2009). This interaction characterizes a spatial phylodynamic process that can be recovered from genomic data using phylogeographic analyses (Faria et al., 2011).

The research activity has been focused on *Pestivirus genus*, that includes pathogens of livestock (Bovine viral diarrhea virus) and wildlife (Border disease virus), and on zoonotic emerging diseases, involving in their epidemiological cycle both livestock and wildlife (Crimean-Congo hemorrhagic fever - CCHF, Hepatitis E). Concerning Pestivirus, several newly emerged genetic variant were detected both in domestic and wild ruminant populations, with mortalities ranging from 40% to 85% in Pyrenean chamois (Marco et., 2015). CCHF is a zoonosis mainly transmitted by ticks that causes sporadic cases and severe hemorrhagic fever of acute human disease with a mortality rate of 5-60% and it has recently emerged in the Balkans and eastern Mediterranean areas (Ergonul, 2006). Hepatitis E virus genotype 3 is the main cause of sporadic acute human hepatitis in Europe (Dalton et al., 2008) and there is evidence indicating that HEV infection can be transmitted by undercooked offal or meat of pigs or wild ungulates (Hoofnagle et al., 2012).

On the basis of above considerations, a phylogenetic approach to analyses the spatial-temporal dynamics of RNA viral sequences belonging to the selected diseases can contribute to reconstruct the most probable places of origin and pathways of dispersion. This approach potentially helps to predict the emergence of infectious diseases by identifying geographic areas from which new infections are likely to emerge and to elucidate the impact of natural or human mediated animal movement or human mobility on viral disease spread (Faria et al., 2011)

CHAPTER 1
PESTIVIRUS IN DOMESTIC AND WILD RUMINANTS

SPATIAL AND TEMPORAL PHYLOGENY OF BORDER DISEASE VIRUS IN PYRENEAN CHAMOIS

ABSTRACT

Border disease virus (BDV) (Genus Pestivirus, Family Flaviviridae) affects a wide range of ruminants worldwide, mainly domestic sheep and goat. Since 2001 several outbreaks of disease associated to BDV infection have been described in Pyrenean chamois (Rupicapra pyrenaica pyrenaica) in Spain, France and Andorra. These outbreaks have decimated several Pyrenean chamois populations, with mortalities ranging from 40% to 85%. The infection has become endemic in the Central and Eastern Pyrenees. After the severe BDV outbreaks, different epidemiological scenarios have appeared in the Pyrenees, some of which are having a negative impact on host population dynamics. The aim of this study was to clarify the origin and dispersion of the Pyrenean chamois BDV genetic variant by reconstructing the spatial and temporal dynamics of BDV 5' UTR sequences of Pyrenean chamois, 10 novel sequences and 41 retrieved from public databases. Sheep BDV sequences (n=43) from Spain and France were also retrieved. A phylogenetic analysis was performed using a Bayesian Markov chain Monte Carlo (MCMC) method implemented in the BEAST v.1.74 package. The maximum clade credibility tree summarizing all of the trees obtained during the MCMC search showed a main clade supported by posterior probabilities of 1, corresponding to the Pyrenean chamois phylogenetic group. The chamois clade originated from sheep BDV and the time of the most recent common ancestor estimates for the chamois clade dated back to a time span between 1974 and 1996, with a mean

estimation falling in 1988,, thus indicating its relatively recent emergence.. There were also some highly significant subclades among chamois sequences clustering three geographical areas in Pyrenees: Eastern area (Freser); Central area (Cadì, Cerdanya-Alt Urgell, Orlu) and Western area (Val d'Aran, Pallars Sobira, Aragon and Andorra). Our data suggest that Pyrenean chamois phylogenetic group originated from sheep BDV genotype 4, generating a founder effect due to intra-species spread and spatial dispersion, still going on such as Western direction.

INTRODUCTION

Pestiviruses, like other RNA viruses, are characterized by a high degree of genetic variability and a high rapidity of evolution, accumulating measurable differences over the time scale of human observation. Genus Pestivirus comprises four approved species, namely Border disease virus (BDV), Bovine viral diarrhoea virus type 1 (BVDV-1) and type 2 (BVDV-2), and Classical swine fever virus (CSFV), traditionally classified according to the host species of isolation. In the last decades, identification and genetic characterization of pestivirus strains in different animal species revealed an extensive interspecies transmission among both domestic (Vilcek & Belak, 1996; Cranwell et al., 2007; Braun et al., 2013) and wild ungulates (Becher et al., 1997; 2003), demonstrating a low host specificity and a wide host range of the genus. Besides the interspecies transmission of known genetic variants, several newly emerged pestivirus were detected in livestock (Hamers et al., 2001; Hurtado et al., 2003; Thabti et al., 2005) and in wild ruminant populations (Marco et al., 2009). Genetic changes of viruses can lead to alterations of virulence and adaptation to new

hosts, as recently observed with the Bungowannah pestivirus causing myocarditis with high mortality in pigs (Kirkland et al., 2007; Finlaison et al., 2009) and the emergence of ovine pestiviruses more closely related to CSFV than to ruminant pestiviruses in Tunisia and Spain (Hurtado et al., 2003; Thabti et al., 2005). Phylogenetic analysis of pestiviruses is therefore crucial for classifying novel viruses and to the study of the evolutionary dynamics, clarifying the relationships between genetic diversity and its temporal-spatial distribution in animal populations.

BDV affects a wide range of ruminants worldwide, mainly domestic sheep and goat and segregates into at least eight phylogenetic groups (Giammarioli et al., 2011), namely BDV-1 to BDV-7. Moreover isolates from Turkey (Oguzoglu et al., 2009) and Tunisia (Thabti et al., 2005) possibly form further groups.

Concerning pestivirus in chamois, since 2001 several outbreaks of disease associated to BDV-4 infection have been described in Pyrenean chamois (Rupicapra pyrenaica pyrenaica) in Spain, France and Andorra (Marco, 2012). These outbreaks have decimated several Pyrenean chamois populations, with mortalities ranging from 40% to 85% (Marco et al., 2015). After the severe BDV outbreaks, different epidemiological scenarios have appeared in the Pyrenees, some of which are having a negative impact on host population dynamics (Fernandez-Sirera et al., 2012a). In respect to BDV in Alpine chamois (Rupicapra rupicapra), a BDV-6, a genetic variant circulating in sheep, has been recently detected in a single animal in French Southern Alps, the chamois population showed a seroprevalence of 38.7% and no mortality or clinical diseases (Martin et al., 2015). Moreover in Switzerland a comprehensive investigation on wild ruminants showed a sporadic frequency of infections and identified a single chamois

BVDV-1h positive (Casaubon et al., 2012), also in this case a genetic variant of livestock origin.

Pestivirus antibody prevalence in other Alpine countries showed that prevalence ranged from absence in Austria (Krametter et al., 2004) and to absence (Citterio et al., 2003) or low prevalence in Italian Central Alps (Gaffuri et al., 2006) and in Italian Western Alps (Olde Riekerink et al., 2005), without outbreaks mortality.

In order to reconstruct the most probable places of origin and pathways of dispersion of the Pyrenean chamois BDV genetic variant, a phylogenetic analysis of BDV 5'untranslated (UTR) sequences of chamois and domestic ungulates (including novel sequences and retrieved from public databases) has been performed by using a new Bayesian framework allowing the spatial–temporal reconstruction of the evolutionary dynamics of highly variable viruses (Lemey et al., 2009).

MATERIALS AND METHODS

Samples and sequence data set

The BDV positive samples and the derived sequences from chamois were selected on the basis of the following criteria: known locality where the strain was isolated and sampling dates. A total of 51 chamois BDV 5'UTR sequences were available, encompassing an alignment of 234 nucleotides, 10 of which were novel Pyrenean chamois sequences and 41 were retrieved from public database. (http://www.ncbi.nlm.nih.gov/nuccore/?term=pestivirus+chamois, accessed last time at 17-3-2015), On the whole 50 sequences were from Pyrenean chamois and one from

Alpine chamois (Rupicapra rupicapra) and the sampling dates ranged from 1996 to 2011.

BDV sequences from domestic ungulates were also retrieved from public database, restricting the geographic localization to Spain and France, which are the Countries of origin of available sequences of chamois. A total of 44 BDV 5'UTR sequences were available, encompassing an alignment of 234 nucleotides, 43 of which were from sheep and one from a pig and the sampling dates ranged from 1984 to 2007 http://www.ncbi.nlm.nih.gov/nuccore/?term=border+disease+virus+france; http://www.ncbi.nlm.nih.gov/nuccore/?term=border+disease+virus+spain, accessed last time at 17-3-15). Origins and characteristics of data set BDV strains are summarized in Table 1 and 2, chamois and domestic ungulates respectively.

Table 1 Accession numbers, references, localities and collection years of the 51 BDV sequences of chamois included in the data set

Name	Host	Accession no.	Reference	Origin	Year
Cfr1@96	R. pyrenaica	FN691777	Marco *et al.*, 2011	Freser-Setcases	1996
Cfr2@96	R. pyrenaica	FN691778	Marco *et al.*, 2011	Freser-Setcases	1996
Can1@02	R. pyrenaica	AY738080	Arnal *et al.*, 2004	Andorra	2002
Cpa1@02	R. pyrenaica	AY641529	Hurtado *et al.*, 2004	Alt Pallars	2002
Cor6@04	R. pyrenaica	EU477593	unpublished	Orlu	2004
Cce1@04	R. pyrenaica	AM905930	Marco *et al.*, 2009	Cerdanya	2004
Cce2@05	R. pyrenaica	AM905931	Marco *et al.*, 2009	Cerdanya	2005
Cce3@05	R. pyrenaica	AM905932	Marco *et al.*, 2009	Cerdanya	2005
Cce4@05	R. pyrenaica	AM905933	Marco *et al.*, 2009	Cerdanya	2005
Cca1@06	R. pyrenaica	AM905918	Cabezon *et al*., 2010	Cadí	2006
Cca2@05	R. pyrenaica	AM905919	Marco *et al.*, 2009	Cadí	2005
Cca3@05	R. pyrenaica	AM905920	Marco *et al.*, 2009	Cadí	2005
Cca4@05	R. pyrenaica	AM905921	Marco *et al.*, 2009	Cadí	2005
Cca5@06	R. pyrenaica	AM905922	Marco *et al.*, 2009	Cadí	2006
Cca6@06	R. pyrenaica	AM905923	Marco *et al.*, 2009	Cadí	2006
Cca7@06	R. pyrenaica	AM905924	Cabezon *et al*., 2010	Cadí	2006
Cca8@06	R. pyrenaica	AM905925	Cabezon *et al*., 2010	Cadí	2006
Cca9@06	R. pyrenaica	AM905926	Cabezon *et al*., 2010	Cadí	2006
Cca10@06	R. pyrenaica	AM905927	Cabezon *et al*., 2010	Cadí	2006
Cca11@06	R. pyrenaica	AM905928	Cabezon *et al*., 2010	Cadí	2006
Cca12@06	R. pyrenaica	AM905929	Cabezon *et al*., 2010	Cadí	2006
Cca13@06	R. pyrenaica	FN397676	Cabezon *et al*., 2010	Cadí	2006
Car1@05	R. pyrenaica	AM765800	Marco *et al.*, 2008	Aran	2005
Car2@05	R. pyrenaica	AM765801	Marco *et al.*, 2008	Aran	2005
Car3@05	R. pyrenaica	AM765802	Marco *et al.*, 2008	Aran	2005
Car4@05	R. pyrenaica	AM765803	Marco *et al.*, 2008	Aran	2005
Car5@06	R. pyrenaica	AM765804	Marco *et al.*, 2008	Aran	2006
Car6@06	R. pyrenaica	AM765805	Marco *et al.*, 2008	Aran	2006
Car7@06	R. pyrenaica	AM765806	Marco *et al.*, 2008	Aran	2006
Car8@06	R. pyrenaica	AM765807	Marco *et al.*, 2008	Aran	2006
Car9@08	R. pyrenaica	HE818617	Fernández-Sirera *et al*., 2012a	Aran	2008
Car10@08	R. pyrenaica	HE818618	Fernández-Sirera *et al*., 2012a	Aran	2008
Car11@11	R. pyrenaica	HE818619	Fernández-Sirera *et al*., 2012a	Aran	2011
Car12@11	R. pyrenaica	HE818620	Fernández-Sirera *et al*., 2012a	Aran	2011
Car13@11	R. pyrenaica	HE818621	Fernández-Sirera *et al*., 2012a	Aran	2011
Car14@11	R. pyrenaica	HE818622	Fernández-Sirera *et al*., 2012a	Aran	2011
Can2@09	R. pyrenaica	HE615083	Fernández-Sirera *et al*., 2012b	Andorra	2009
Can3@09	R. pyrenaica	HE615084	Fernández-Sirera *et al*., 2012b	Andorra	2009
Can4@09	R. pyrenaica	HE615085	Fernández-Sirera *et al*., 2012b	Andorra	2009
C4606@06	R. pyrenaica	EU637005	unpublished	Ariege	2006
CFSA1@11	R. rupicapra	KC859383	Martin *et al*., 2015	French Alpes	2011
Cri1@09	R. pyrenaica	-	unpublished	Alta Ribagorça	2009
Cri2@09	R. pyrenaica	-	unpublished	Alta Ribagorça	2009

Cri3@09	R. pyrenaica	-	unpublished	Alta Ribagorça	2009
Cri4@09	R. pyrenaica	-	unpublished	Alta Ribagorça	2009
Cpa2@02	R. pyrenaica	-	unpublished	Alt Pallars	2002
Cpa3@02	R. pyrenaica	-	unpublished	Alt Pallars	2002
Cpa4@02	R. pyrenaica	-	unpublished	Alt Pallars	2002
Cpa5@02	R. pyrenaica	-	unpublished	Alt Pallars	2002
Cpa6@02	R. pyrenaica	-	unpublished	Alt Pallars	2002

Table 2 Accession numbers, references, localities and collection years of the 44 BDV sequences of sheep and pig included in the data set

Name	Host	Accession no.	Reference	Origin	Year
Oll1@97	sheep	FR714860	unpublished	Spain Lleida-Catalunya	1997
Oll2@97	sheep	HE999869	unpublished	Spain Lleida-Catalunya	1997
Oll3@97	sheep	HE999870	unpublished	Spain Lleida-Catalunya	1997
Oll4@97	sheep	HE999871	unpublished	Spain Lleida-Catalunya	1997
Oll5@97	sheep	HE999872	unpublished	Spain Lleida-Catalunya	1997
Oal1@99	sheep	AY159513	Hurtado *et al.*, 2003	Spain Alava	1999
Oal2@01	sheep	AY159516	Hurtado *et al.*, 2003	Spain Alava	2001
Oal3@01	sheep	AY159517	Hurtado *et al.*, 2003	Spain Alava	2001
Oal4@01	sheep	AY159515	Hurtado *et al.*, 2003	Spain Alava	2001
Ote1@04	sheep	DQ275625	Valdazo-Gonzalez *et al.*, 2007	Spain Teruel	2004
Ote2@04	sheep	DQ275623	Valdazo-Gonzalez *et al.*, 2007	Spain Teruel	2004
Ote3@04	sheep	DQ275624	Valdazo-Gonzalez *et al.*, 2007	Spain Teruel	2004
Ole1@01	sheep	DQ275626	Valdazo-Gonzalez *et al.*, 2007	Spain Leon	2001
Ole2@02	sheep	DQ361071	Valdazo-Gonzalez *et al.*, 2006	Spain Leon	2002
Ole3@01	sheep	DQ361072	Valdazo-Gonzalez *et al.*, 2006	Spain Leon	2001
Obu1@02	sheep	DQ275622	Valdazo-Gonzalez *et al.*, 2007	Spain Burgos	2002
Obu2@02	sheep	DQ361069	Valdazo-Gonzalez *et al.*, 2006	Spain Burgos	2002
Obu3@02	sheep	DQ361067	Valdazo-Gonzalez *et al.*, 2006	Spain Burgos	2002
Obu4@02	sheep	DQ361068	Valdazo-Gonzalez *et al.*, 2006	Spain Burgos	2002
Oza1@02	sheep	DQ361070	Valdazo-Gonzalez *et al.*, 2006	Spain Zamora	2002
Ocu1@01	sheep	DQ361073	Valdazo-Gonzalez *et al.*, 2006	Spain Cuenca	2001
Pcl1@07	pig	HF567456	Rosell *et al.*, 2014	Spain Castilla Leon	2007
Opa1@06	sheep	EF694003	Dubois *et al.*, 2008	France PACA	2006
Opa2@06	sheep	EF694002	Dubois *et al.*, 2008	France PACA	2006
Opa3@06	sheep	EF694001	Dubois *et al.*, 2008	France PACA	2006
Opa4@06	sheep	EF694000	Dubois *et al.*, 2008	France PACA	2006
Opa5@06	sheep	EF693999	Dubois *et al.*, 2008	France PACA	2006
Omp1@96	sheep	EF693998	Dubois *et al.*, 2008	France Midi-Pyrénées	1996
Opa6@94	sheep	EF693997	Dubois *et al.*, 2008	France PACA	1994
Opa7@94	sheep	EF693996	Dubois *et al.*, 2008	France PACA	1994
Opa8@93	sheep	EF693995	Dubois *et al.*, 2008	France PACA	1993
Opa9@92	sheep	EF693994	Dubois *et al.*, 2008	France PACA	1992
Opa10@91	sheep	EF693993	Dubois *et al.*, 2008	France PACA	1991
Opa11@90	sheep	EF693992	Dubois *et al.*, 2008	France PACA	1990
Opa12@90	sheep	EF693991	Dubois *et al.*, 2008	France PACA	1990
Opa13@90	sheep	EF693990	Dubois *et al.*, 2008	France PACA	1990
Opa14@90	sheep	EF693989	Dubois *et al.*, 2008	France PACA	1990
Oaq1@89	sheep	EF693988	Dubois *et al.*, 2008	France Aquitaine	1989
Opa15@89	sheep	EF693987	Dubois *et al.*, 2008	France PACA	1989
Opa16@85	sheep	EF693986	Dubois *et al.*, 2008	France PACA	1985
Oce1@85	sheep	EF693985	Dubois *et al.*, 2008	France Centre	1985
OAV1@84	sheep	EF693984	Dubois *et al.*, 2008	France Midi-Pyrénées	1984
Opa17@91	sheep	EF988632	Dubois *et al.*, 2008	France PACA	1991
Opa18@91	sheep	EF988633	Dubois *et al.*, 2008	France PACA	1991

RT-PCR and sequencing

Viral RNA of the 10 novel sequences was extracted from chamois spleen tissues using a commercial RNA extraction kit (Nucleospin Viral RNA Isolation, Macherey Nagel, Germany). Reverse transcription and PCR assays targeting 5'UTR region of pestivirus were performed using previously described 324 and 326 pan-pestivirus primers (Vilcek et al., 1994; 1996) following previously described protocols (Fernández-Sirera et al., 2012). For each sample, the amplicons of the expected sizes were purified and sequenced using forward and reverse primers by cycle sequencing using a Big Dye Terminator version 3.1 kit and an ABI PRISM 3130xl sequencing device (Applied Biosystems, CA, USA). The sequences will be submitted to GenBank.

All the chamois and sheep available BDV sequences were aligned with BDV reference strains, retrieved from GenBank and representative of BDV phylogenetic groups according to Giammarioli et al., (2011), using CLUSTALW (integrated within the Bio-Edit sequence editor by Tom Hall, 2001; http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Phylogeny was estimated by the neighbor-joining algorithm (NJ) and the maximum likelihood (ML) method.

Likelihood mapping

The phylogenetic signal of each sequence dataset was investigated by means of the likelihood-mapping analysis of 10,000 random quartets generated using TreePuzzle. All of the three possible unrooted trees for a set of four sequences (quartets) randomly selected from the dataset were reconstructed using the maximum likelihood approach and the selected substitution model. The posterior probabilities of each tree were then

plotted on a triangular surface so that the dots representing the fully resolved trees fell at the corners and those representing the unresolved quartets in the centre of the triangle (star-tree) (Schmidt et al 2002). Using this strategy, which has been described in detail elsewhere (Strimmer & von Haeseler 1997), the data are considered unreliable for phylogenetic inference when more than 30% of the dots are in the centre of the triangle.

Phylogenetic reconstruction

The best-fitting nucleotide substitution model was estimated by means of JModeltest (Posada, 2008), and selected a TrN model (Rodriguez et al., 1990) with gamma-distributed rates among sites.

The phylogenetic tree, model parameters, evolutionary rates and population growth were co-estimated using a Bayesian Markov chain Monte Carlo (MCMC) method implemented in the BEAST v.1.74 package (Drummond et al., 2012).

Statistical support for specific clades was obtained by calculating the posterior probability of each monophyletic clade. As coalescent priors, we compared four simple parametric demographic models (constant population size, and exponential, expansion and logistic population growth) and a piecewise-constant model, the Bayesian skyline plot (BSP) under both a strict and a relaxed (uncorrelated log-normal) clock (Drummond et al, 2005)..

Two independent MCMC chains were run for 50 million generations with sampling every 5,000th generation, and were combined using the LogCombiner 1.74 included in the BEAST package. Convergence was assessed on the basis of the effective sampling

size (ESS) after a 10% burn-in using Tracer software version 1.5 (http://tree.bio.ed.ac.uk/software/tracer/). Only ESS's of ≥ 200 were accepted.

Uncertainty in the estimates was indicated by 95% highest posterior density (95% HPD) intervals, and the best-fitting models were selected using a Bayes factor (BF with using marginal likelihoods) implemented in BEAST (Suchard et al., 2001).

In accordance with Kass & Raftery (1995), only values of 2lnBF ≥ 6 were considered significant.

The trees were summarised in a target tree using the Tree Annotator program included in the BEAST package, choosing the tree with the maximum product of posterior probabilities (maximum clade credibility) after a 10% burn-in.

The time of the most recent common ancestor (tMRCA) estimates were expressed as mean and 95% HPD years before the most recent sampling dates, corresponding to 2011 in this study.

Bayesian phylogeography

The continuous-time Markov Chain (MCC) process over discrete sampling locations implemented in BEAST (Drummond & Rambaut, 2007) was used for the geographical analysis, implementing the Bayesian Stochastic Search Variable Selection (BSSVS) model which allows the diffusion rates to be zero with a positive prior probability. Comparison of the posterior and prior probabilities of the individual rates being zero provided a formal BF for testing the significance of the linkage between locations. This analysis was performed for Pyrenean chamois dataset (n=50).

The final tree was visualised using FigTree version 1.4 (available at http://tree.bio.ed.ac.uk/software). The significant migration rates were analysed and

visualised using SPREAD, which is available at http://www.kuleuven.be/aidslab/phylogeography/SPREAD.html. The 50 isolates were assigned to a total of 9 distinct geographic groups corresponding to: Alt Pallars (n=7), Alta Ribagorça (n=4), Andorra (n=4), Aran (n=14), Cadì (n=13), Cerdanya-Alt Urgell (n=4), Orlu (n=1), Freser-Setcases (n=2), unspecified area of france Pyrenees (n=1).

In order to provide a spatial projection, the migration routes indicated by the tree were visualised using Google Earth (http://earth.google.com).

RESULTS

Likelihood mapping

The likelihood mapping of 10,000 random quartets showed that more than 75% were distributed at the corners of the likelihood map and only 24.3% in the central area, thus indicating that the dataset contained sufficient phylogenetic information in supplementary Fig. S1 (see Annex section).

Bayesian phylogeny and phylogeography

Fig. 1 represents the maximum clade credibility tree summarizing all of the trees obtained during the MCMC search on BDV sequences from domestic ungulates and chamois in France, Andorra and Spain. The ovine sequences clustered in six main clades, supported by posterior probabilities ≥ 0.99, corresponding to the known viral genetic groups BDV-3, BDV-4a, 4b, BDV-5, BVD-6 and Tunisian BDV. The single pig sequence belongs to BDV-4a. Pyrenean chamois sequences clustered in a unique clade, supported by posterior probability of 1, that originated from BDV-4a sheep clade (pp=0.86). The mean tMRCA of BDV-4a and the Pyrenean chamois clades existed a

mean 24 years ago (95%HPD 17-35 years ago).There were also some highly significant subclades (pp=1) among Pyrenean chamois sequences clustering three geographical areas in Pyrenees: Eastern area (Freser-Setcases), Central area (Cadí, Cerdanya-Alt Urgell, Orlu) and Western area (Alt Pallars, Alta Ribagorça, Aran and Andorra). The single French Alpine chamois sequence clustered in BDV 6 ovine genetic group as recently reported (Martin et al., 2015).

Fig. 1 The maximum clade credibility tree of BDV sequences from domestic ungulates and chamois in France, Andorra and Spain. The numbers on the internal nodes indicate significant posterior probabilities (pp>0.85), and the scale at the bottom of the tree represents the number of years before the last sampling time (2011). The chamois main geographical clades and BDV subtypes have been highlighted.

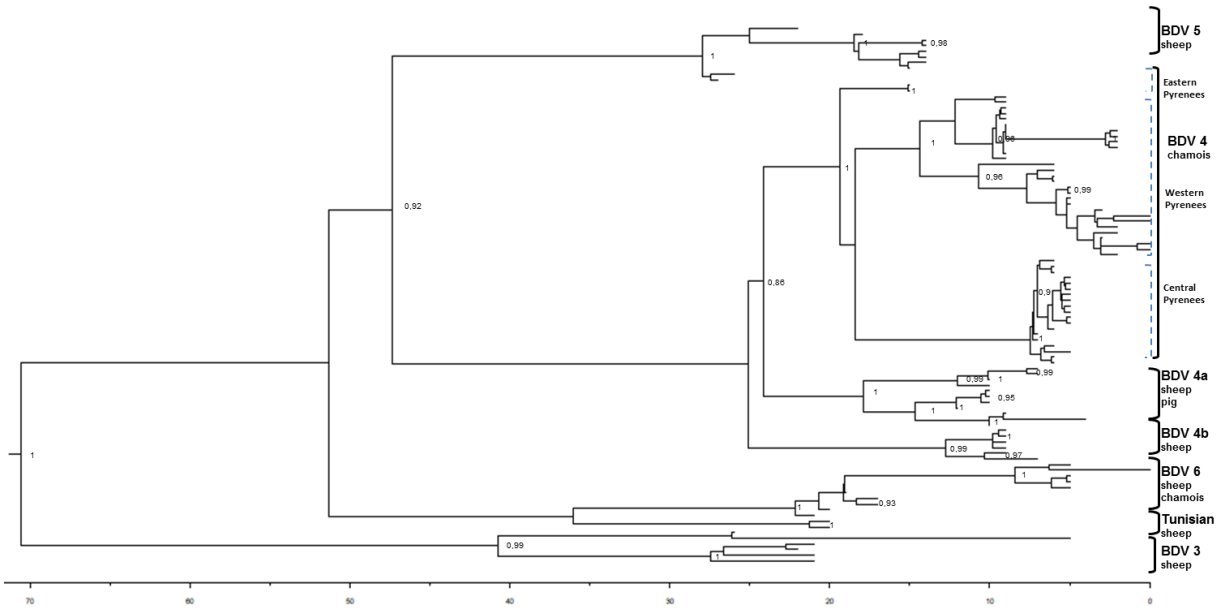


Figure 2 represents the phylogeographic MCC tree, obtained grouping only Pyrenean chamois sequences (n=50) within nine groups on the basis of the sampling location (AND, Andorra; APA, Alt Pallars; ARA, Aran; ARI, Alta Ribagorça; CAD, Cadí; CER, Cerdanya-Alt Urgell; FRE, Freser-Setcases; ORL, Orlu; FRA, unspecified area of French Pyrenees). The analysis of the tree highlighted a significant geographic structure, the strains obtained in one locality tended to segregate significantly from those isolates in different area. Only strains from Andorra are interspersed within Alt Pallars and Aran clades. There were three main clades, as already observed in figure 1 and some significant subclades. The most probable location of the root of the tree was Freser-Setcases, supported by the highest state posterior probability (pp=0.39 vs pp=0.09 for Andorra, the second most probable location of the root). The tMRCA for the tree root was estimated to be 22.95 years (95% HPD= 15-37.7), corresponding to 1988.

Figure 2 The MCC tree maximum clade credibility (MCC) tree of BDV-4 5'UTR sequences of Pyrenean chamois sequences (n=50). The branches are coloured on the basis of the most probable location of the descendent nodes (AND, Andorra; APA, Alt Pallars; ARA, Aran; ARI, Alta Ribagorça; CAD, Cadí; CER, Cerdanya-Alt Urgell; FRE, Freser-Setcases; ORL, Orlu; FRA, unspecified area of French Pyrenees).

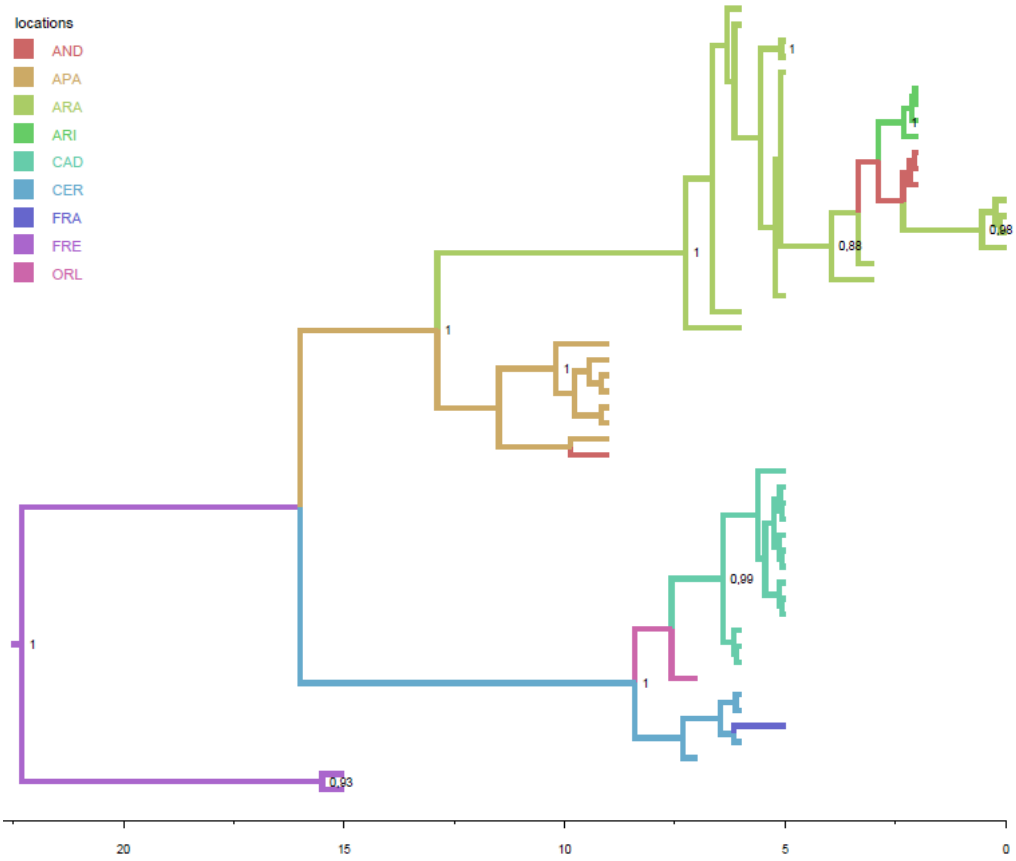


Table 3 summarized the tMRCA of the root and internal nodes and the most probable locations for different clades and subclades. In particular the MRCA shared by the subclades of Western Pyrenees was located in Alt Pallars (pp=0.53 vs 0.08 for Aran) and the mean tMRCA was 13.45. The MRCA of Eastern Pyrenees was located in Cerdanya-Alt Urgell (pp=0.42 vs 0.11 for Cadì) and the tMRCA was 8.28 years.

Bayesian phylogeography showed statistically well supported links at the Bayes factor

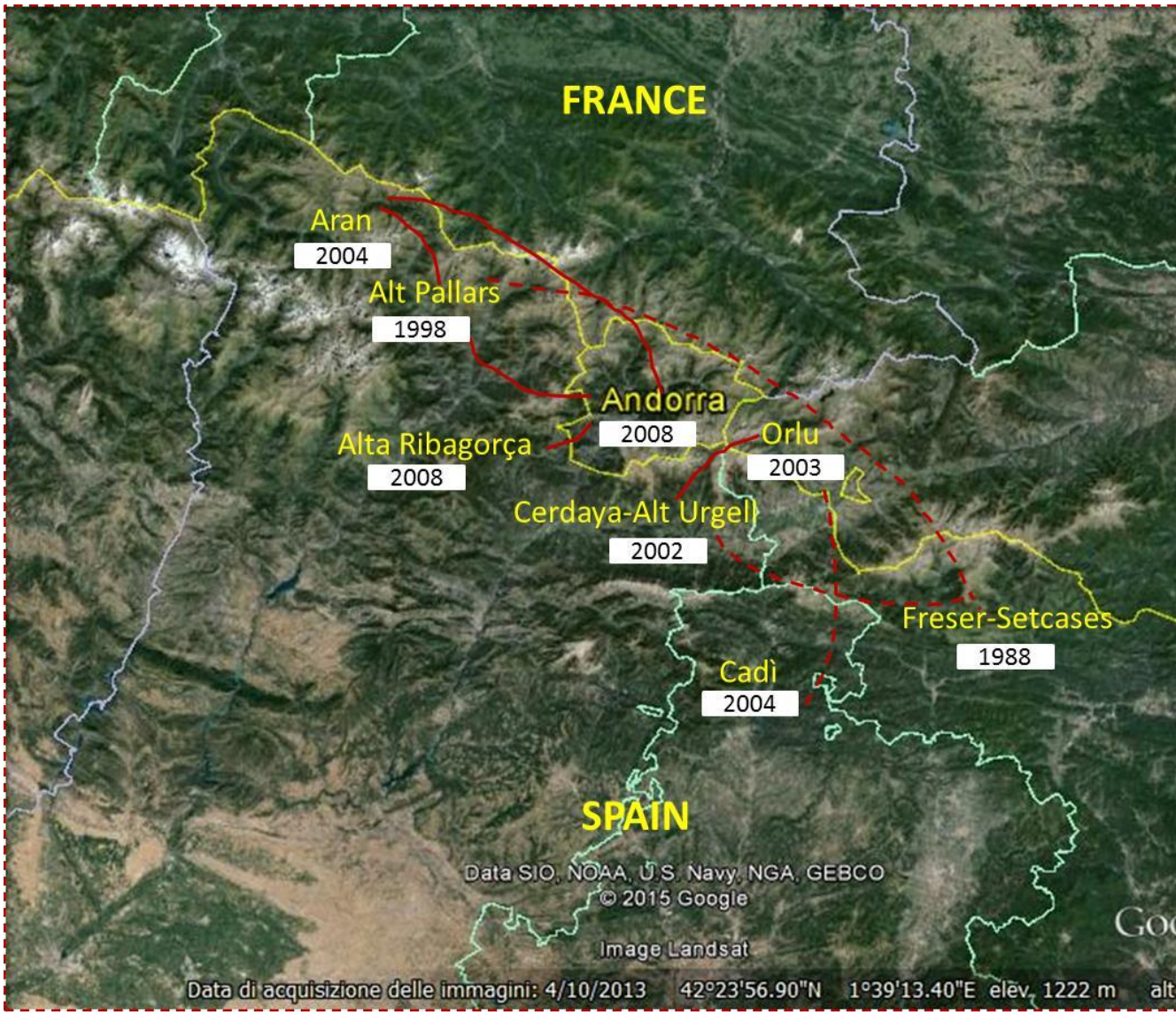
test (BF >3) between the following geographic localities: Alt Pallars and Andorra, Aran and Andorra, Andorra and Alta Ribagorça , Cerdanya-Alt Urgell and Orlu. Figure 3 summarized the dispersion patterns of Pyrenean chamois BDV sequences as reconstructed on the basis of significant genetic flow rates and dated phylogeny.

Table 3 Time of the most common ancestor estimates of Pyrenean chamois BDV, credibility interval (95%HPD) of the main clades observed in the MCC tree, with the corresponding most probable locations, and state posterior probability

CLADE	Subclade	tMRCA[1]	CI tMRCA L[2]	CI tMRCA U[3]	Locality	State pp[4]
ROOT		23	15,0	37,7	FRESER-SETCASES	0,39
EASTERN PYRENEES		15,7	15,0	17,4	FRESER-SETCASES	0,96
WESTERN PYRENEES		13,5	9,6	19,2	ALT PALLARS	0,53
	ALT PALLARS	12	9,3	15,7	ALT PALLARS	0,73
	ARAN	7,7	6,2	9,7	ARAN	0,92
CENTRAL PYRENEES		8,3	7,0	9,9	CERDANYA	0,42
	CADÍ	6,5	6,0	7,3	CADÍ	0,97

[1] tMRCa: time of the most common ancestor
[2] CI tMRCA L: lower credibility interval
[3] CI tMRCA U: upper credibility interval
[4] pp: posterior probability

Figure 3 Significant non-zero BDV-4 Pyrenean chamois migration rates in Pyrenees areas. Sstatistically well supported links at the Bayes factor test (BF >3) are shown in red continuous line.



DISCUSSION

Identification and genetic characterization of BDV strains from Pyrenean chamois have been performed since the first outbreaks, indeed the intensive monitoring of hunted or found dead chamois allowed to collect a large number of strains from different areas in the Pyrenees (Marco et al., 2009) and therefore to apply advanced phylogenetic analysis. Previous investigations performed phylogenetic analysis using the neighbor-joining (NJ) method and classified Pyrenean chamois strains within BDV-4 genotype (Arnal et al., 2004; Hurtado et al., 2004).

In order to reconstruct the origin, time of introduction and pathways of dispersion of the Pyrenean chamois BDV genetic variants, we performed a phylogenetic analysis of a comprehensive collection of publicly available ovine and chamois BDV sequences (including novel Pyrenean chamois sequences) of Spanish, Andorran and French origin by using a new Bayesian framework allowing the spatial–temporal reconstruction of the evolutionary dynamics of highly variable viruses (Lemey et al., 2009). A total of 95 5'UTR sequences were included, 51 out of 95 of chamois origin and collected over a time span of 16 years (1996 to 2011).

All the Pyrenean chamois isolates cluster in a unique highly significant clade within the BDV-4 genotypes, including sequences from Spain, France and Andorra, in agreement with previous analysis using NJ methods on more restricted datasets (Arnal et al., 2004; Hurtado et al., 2004). On the whole BDV-4 includes also Spanish ovine isolates and segregate into three highly significant subclades: BDV-4a, BDV-4b (Valdazo-Gonzalez et al., 2007) and a third one of the Pyrenean chamois isolates. The

phylodynamic and phylogeographical analysis suggests that BDV-4a and the Pyrenean chamois clades shared a common ancestor and therefore that the chamois viral genetic variant originated from sheep. The entire diversity of Pyrenean chamois genetic group is likely descended from a single introduction that generated a founder effect within the chamois population due to intra-species spread and spatial dispersion in different Pyrenean areas.

Moreover, the Pyrenean chamois sequences showed a strong spatial structure and clustered in three geographical areas within the Pyrenees: Eastern area (Freser-Setcases), Central area (Cadí, Cerdanya-Alt Urgell, Orlu) and Western area (Alt Pallars, Alta Ribagorça, Aran and Andorra).

In our reconstruction, BDV chamois genetic variant originated in Freser-Setcases dating back to a time span between 1974 and 1996, with a mean estimation falling in 1988. This period of time confirms the relatively recent emergence of BDV-4 in Pyrenean chamois and is highly consistent with retrospective sero-epidemiological and molecular investigations that reported BDV at least since 1990 and 1995 respectively in Spanish Pyrenees (Marco et al., 2011) and in French Pyrenees (Pioz et al., 2007). Two highly divergent Pyrenean chamois clades originated from Freser-Setcases: the Western Pyrenean and the Central Pyrenean dating back to 1998 and 2003 respectively as a mean estimation. The first one gave origin to Alt Pallars outbreak and subsequently spreading to neighboring areas of Aran, Andorra and Alta Ribagorça, that resulted significantly linked by viral migration flows until recently. The pathway of dispersion of isolates among these areas suggests a complex exchange between neighboring areas still going on, supporting the epidemiological evidence of a frequent

circulation of BDV in Western Pyrenees (Fernández-Sirera et al., 2012a). Interestingly, Alt Pallars clade included all the isolates from the BDV outbreak occurred in 2001 and 2002 (Marco et al., 2007), except a single isolate from Andorra, reported in 2002 by (Arnal et al., 2004), finally claryfing that this strain is not related to the Andorra BDV outbreak reported in 2009-2010 (Fernández-Sirera et al., 2012b).

The Central Pyrenean clade, the second main one originating from Freser-Setcases, originally dispersed in Cerdanya-Alt Urgell area, that had an important role in spreading the virus throughout the French Pyrenees and Cadí. Interestingly significant viral migration flows have been observed between Cerdanya- Alt Urgell and French Pyrenees, indicating transmission events between the two chamois populations.

In conclusion our data suggest that Pyrenean chamois phylogenetic group originated from sheep BDV genotype 4, generating a founder effect due to intra-species spread and spatial dispersion, still going on such as Western direction in Pyrenees chamois population. A key factor influencing disease emergence may be pathogen invasiveness through viral mutation (Marco et al., 2015), indeed in other alpine chamois population no mortality outbreaks caused by pestivirus infections have been reported despite interaction with domestic ruminant populations (Citterio et al., 2003; Casaubon et al., 2012, Martin et al., 2015).

EXTENDED GENETIC DIVERSITY OF BOVINE VIRAL DIARRHEA VIRUS AND FREQUENCY OF GENOTYPES AND SUBTYPES IN CATTLE IN ITALY BETWEEN 1995 AND 2013

Biomed Research International 2014:147145. doi: 10.1155/2014/147145.

ABSTRACT

Genetic typing of bovine viral diarrhea virus (BVDV) has distinguished BVDV-1 and BVDV-2 species and an emerging putative third species (HoBi-like virus), recently detected in southern Italy, signaling the occurrence of natural infection in Europe. Recognizing the need to update the data on BVDV genetic variability in Italy for mounting local and European alerts, a wide collection of UTR sequences (n = 371) was selected to identify the frequency of genotypes and subtypes at the herd level. BVDV-1 had the highest frequency, followed by sporadic BVDV-2. No novel HoBi-like viruses were identified. Four distribution patterns of BVDV-1 subtypes were observed: highly prevalent subtypes with a wide temporal-spatial distribution (1b and 1e), low prevalent subtypes with a widespread geographic distribution (1a, 1d, 1g, 1h, and 1k) or a restricted geographic distribution (1f), and sporadic subtypes detected only in single herds (1c, 1j, and 1l). BVDV-1c, k, and l are reported for the first time in Italy. A unique genetic variant was detected in the majority of herds, but cocirculation of genetic variants was also observed. Northern Italy ranked first for BVDV introduction, prevalence, and dispersion. Nevertheless, the presence of sporadic variants in other restricted areas suggests the risk of different routes of BVDV introduction.

INTRODUCTION

Bovine viral diarrhea virus (BVDV), a widespread pathogen of cattle, was first described in North America in 1946 [1, 2]. Seroprevalence rates between 36% and 88% have subsequently been reported since the 1960s in North America, Europe, Australia, and East Africa. The endemic diffusion of BVDV persists in geographic areas where no systematic control measures have been implemented and it reflects the pathogenic mechanisms of BVDV through which the virus can establish both transient and persistent infections. Persistent infected (PI) animals, originating from a transient infection of pregnant cows or born from PI cows, shed large amounts of virus throughout their lives and ensure viral persistence in the host population. Besides the previous maintenance strategy, a peculiar biological characteristic of BVDV contributes to iatrogenic diffusion of infection. In fact two different biotypes coexist and are differentiated by their effect on cell culture in cytopathic (cp) and non-cytopathic biotypes (ncp). Contamination of fetal bovine serum (FBS) by the ncp biotype has long been known and remains a recognized risk factor for the worldwide distribution of BVDV [3, 4]. Because FBS is used in the production of vaccines and other biological products, the global trade of infected FBS products is a potential source of transboundary spread of BVDV.

BVDV belongs to the *Pestivirus* genus of the *Flaviviridae* family. Genetic typing of BVDV isolates distinguishes two recognized species: BVDV-1 and BVDV-2. To date, 17 potential BVDV-1 subtypes have been identified [5, 6, 7, 8, 9, 10] and BVDV-2 strains can be clustered into at least three subtypes [5, 11, 12]: BVDV-1a to BVDV-1q and

BVDV-2a to BVDV-2c, respectively. A putative third bovine species, referred to as HoBi-like virus or BVDV-3 [13], comprises viral strains recently detected in FBS batches originating mainly from South America, as well as from Australia, Canada, Mexico, and the United States [14] . Unlike the widespread diffusion of HoBi-like viruses via FBS, natural infection in cattle has been reported in Southeast Asia [15], South America [16, 17], and in two dairy herds in southern Italy [18, 19]. The Italian cases represent the first identification of HoBi-like virus in cattle in Europe.

In addition to the risk of transboundary spread of BVDV through potentially infected FBS, the genetic diversity of the virus in a given geographic area has been largely influenced by animal movement within countries and/or introduction from other countries, as recently observed in Switzerland [20], Italy [21], and England and Wales [22], showing that the introduction and spatial distribution of BVDV can also be influenced by livestock management practices.

Nucleotide sequencing is a rapid and inexpensive diagnostic tool for unambiguous typing of all bovine pestiviruses. Broad systematic surveillance of BVDV genetic variability is advisable for updating data on the distribution and frequency of emerging genotypes and subtypes in BVDV endemic countries. The aim of this study was therefore to analyze a representative and epidemiologically well-characterized collection of BVDV sequences from Italian cattle. The previous genotyping studies were carried out on a limited number of isolates [23, 24, 25, 26, 27, 28]. The present study represents a comprehensive collection of Italian isolates obtained from all cattle breeding areas over a time period spanning nearly two decades (1995-2013). Genetic

variability was determined to identify the frequency of genotypes and subtypes with resolution at the herd level.

MATERIALS AND METHODS

Samples and data set

The material comprised samples sent to laboratories between 1995 and 2013 for routine testing because of suspected BVDV infection and for screening in voluntary herd control programs. The BVDV positive samples and the derived sequences were selected on the basis of the following criteria: 1) sequences obtained from different animals; 2) known herd and locality where the strain was isolated and sampling dates; and 3) sequences representative of all Italian BVDV subtypes and genotypes. A total of 371 BVDV 5'UTR sequences were included, 272 of which were novel sequences and 99 were BVDV Italian sequences retrieved from published peer-reviewed journals. Samples were obtained from 357 cattle and 14 bulk milk specimens collected from 259 Italian cattle herds from around the country. Samples were obtained from 164 dairy herds, 40 beef herds, 11 mixed production systems, and 44 were undetermined. Only one sequence from each herd was available for the majority of the herds (n=210); 2 to 20 sequences were included for the 49 remaining herds. The localities of origin of the strains were grouped into four macroareas: the North, the Center, the South, and the Islands (Sicily and Sardinia). In detail, the North macroarea comprised the regions of Emilia Romagna (n=41), Lombardy (n=108), Piedmont (n=130), Valle d'Aosta (n=1), and Veneto (n=18); the Center, Latium (n=6), Marches (n=4), Tuscany (n=2), and Umbria

(n=15); the South, Basilicata (n=4), Calabria (n=5), Campania (n=4) and Puglia (n=2); the islands, Sicily (n=29) and Sardinia (n=2).

RT-PCR and sequencing

Viral RNA of the 272 novel sequences was extracted from original biological samples (n=261) and growth medium after passage in cell culture (n=11). Reverse transcription and PCR assays targeting a 288 bp region of 5'UTR of BVDV were performed using previously described primers by Letellier *et al*. [29], with the exception of strains collected in Sicily, which were tested using primers by Vilcek *et al*. [30]. The samples collected in the Center and the South macroareas were also tested by primers for atypical pestivirus [14]. For the BVDV-1 subtypes identified for the first time in Italy, a 428 bp region encoding autoprotease N^{pro} was amplified using nested PCR, as previously described [24, 27].

For each sample, the amplicons of the expected sizes were purified and sequenced using forward and reverse primers by cycle sequencing using a Big Dye Terminator version 1.1 Cycle Sequencing kit (Applied Biosystems, CA, USA) and an ABI PRISM 3130 sequencing device or sent for outsource sequencing (Primm).

Phylogenetic analysis

All the sequences were aligned with BVDV reference strains retrieved from GenBank representative of BVDV-1, BVDV-2, and HoBi-like virus using Clustal X; manual editing was performed with Bioedit software version 7.0 (freely available at http://www.mbio.ncsu.edu/bioedit/bioedit.html). Phylogeny was estimated by the neighbor-joining algorithm (NJ) and the maximum likelihood (ML) method.

The evolutionary model that best fitted the data (GTR + I + G) was selected using an information criterion implemented in JmodelTest [31] (freely available at http://darwin.uvigo.es/software/jmodeltest.html).

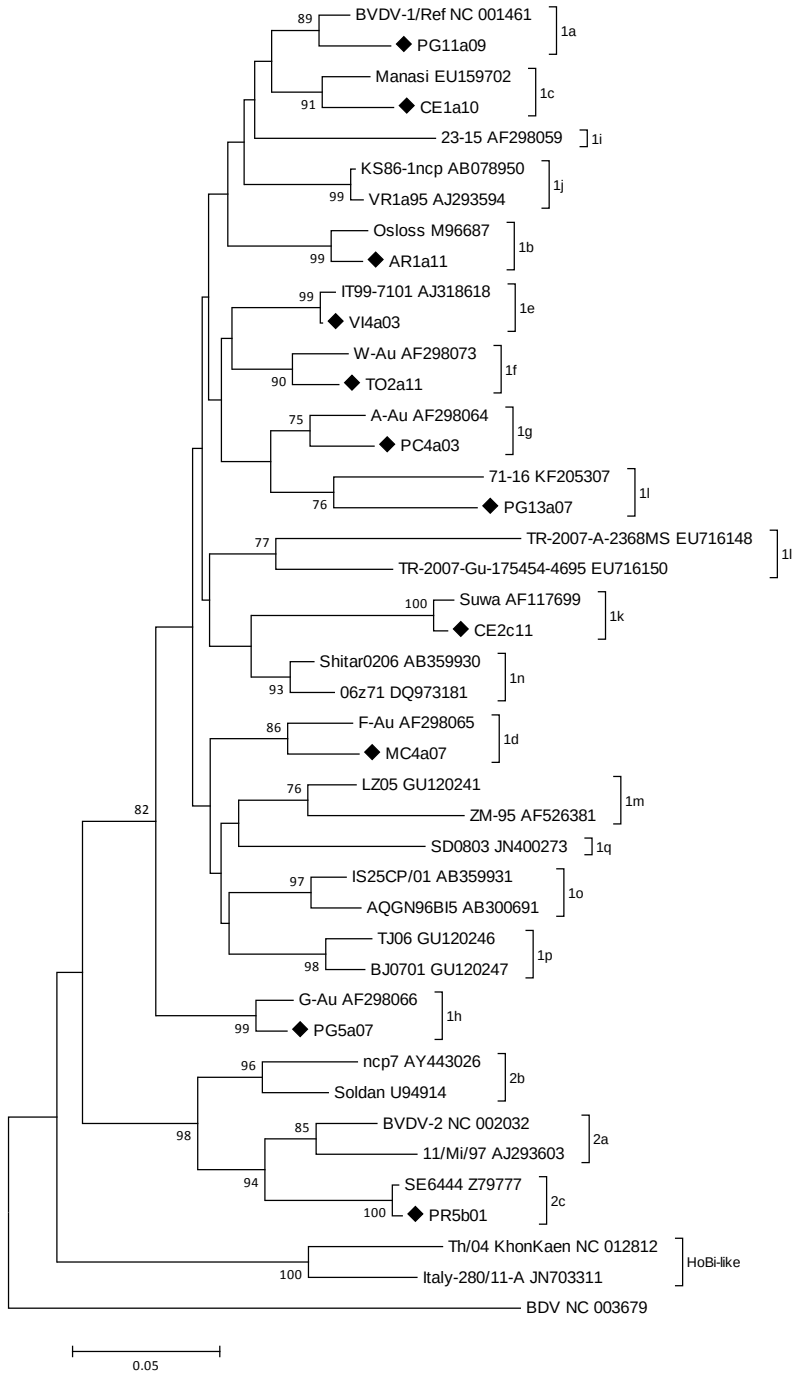
NJ analysis of the 5'UTR was performed using molecular evolutionary genetics analysis (MEGA version 5) [32], with 1000 bootstrap replicates; ML tree was reconstructed with Phyml version 3.0 (http://www.atgc-montpellier.fr/phyml) with 1000 bootstrap replicates.

The sequences within the same genotypes and subtypes from the same herd and from different herds were compared, and the percentage of nucleotide similarity of pairwise evolutionary distances was calculated using MEGA version 5.

RESULTS

A total of 269 out of 272 sequences obtained in the present study were typed as BVDV-1 upon analysis of 5'UTR to reference strains and three were classified as BVDV-2, using both NJ and ML methods. No HoBi-like viruses were identified in the Center and South macroareas using primers described by Xia *et al*. [14] and Letellier *et al*. [29], in northern Italy and Sardinia using primers by Letellier *et al.*[29] and in Sicily using primers described by Vilcek et al. [30]. A selection of Italian BVDV sequences representative of all BVDV genotypes and subtypes detected were reported in the NJ phylogenetic tree of the 5'UTR region, together with reference strains (Figure 1).

Figure 1. Phylogenetic tree based on the 5'-UTR of selected Italian sequences representative of all BVDV genotypes and subtypes detected between 1995 and 2013 and reference BVDV-1, BVDV-2, and HoBi-like strains. Molecular evolutionary genetics analyses were performed with MEGA5 using the NJ method. Distances were computed using the Kimura 2-parameter model. Bootstrap values > 70% are shown. Published sequences and references are identified by GenBank accession number (available at http://www.ncbi.nlm.nih.gov/pubmed/). The symbol "♦" indicates selected novel nucleotide sequences of BVDV Italian strains.



A phylogenetic tree of all the sequences were also represented (Supplementary Figure 1, see Annex Section).

With regard to both the novel and the published sequences, 357 (96.2%) sequences belonged to BVDV-1, ten (2.7%) to BVDV-2, and four (1.1%) to HoBi-like virus. The BVDV-1 sequences belonged to 11 different subtypes (a, b, c, d, e, f, g, h, j, k and l). The frequency of genotypes and subtypes is summarized in Table 1.

Among the subtypes identified, BVDV-1c, k, and l are reported here for the first time in Italy; typing of these subtypes was confirmed by N^{pro} region analysis (data not shown). Phylogenetic analysis showed that the BVDV-2 strains clustered with the subtypes a and c according to [12].

BVDV-1b and BVDV-1e showed the highest frequency at both the animal and the herd levels, being detected in 43.9% and 27.5% of herds, respectively. The frequency of BVDV-1a, d, f, g, h, k was less than 10%; BVDV-1c, j, l were sporadically obtained from single herds. The frequency of BVDV-2 was low at both the animal and the herd levels (2.7% at the herd level, Table 1).

Table 1. Frequency of BVDV genotypes and subtypes in Italy.

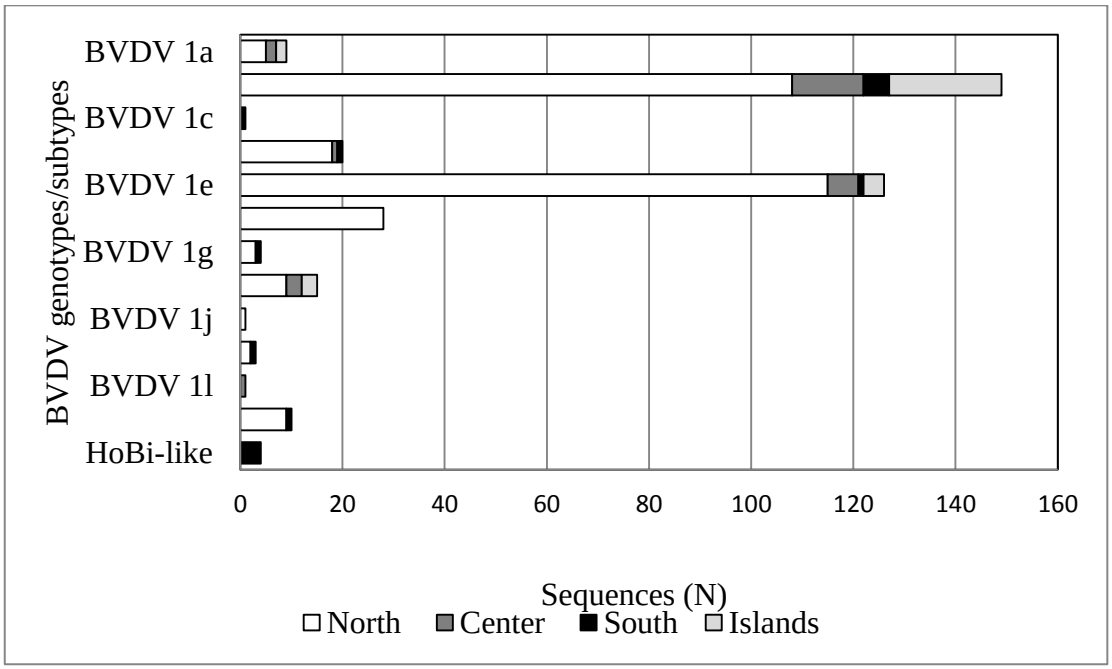
Genotype/subtype	Total sequences No. (%)	Herd* No. (%)	Years	Geographic origin**	Published/ total sequences (n)	References
BVDV-1a	9 (2.4)	9 (3.4)	2000-2009	NCI	3/9	[27]
BVDV-1b	149 (40.2)	115 (43.9)	1995-2013	NCSI	52/149	[21, 23, 25, 27, 28]
BVDV-1c	1 (0.3)	1 (0.4)	2010	S	0/1	Not published
BVDV-1d	20 (5.4)	16 (6.1)	1995-2010	NCS	5/20	[23, 25]
BVDV-1e	126 (34)	72 (27.5)	1996-2013	NCSI	22/126	[21, 23, 25, 27, 28]
BVDV-1f	28 (7.5)	17 (6.5)	1999-2012	N	1/28	[25]
BVDV-1g	4 (1.1)	4 (1.5)	2002-2010	NS	1/4	[27]
BVDV-1h	15 (4.0)	14 (5.3)	1996-2011	NCI	3/15	[21, 23, 25]
BVDV-1j	1 (0.3)	1 (0.4)	1995	N	1/1	[23]
BVDV-1k	3 (0.8)	3 (1.1)	2001-2011	NS	0/3	Not published
BVDV-1l	1 (0.3)	1 (0.4)	2007	C	0/1	Not published
BVDV-2	10 (2.7)	7 (2.7)	1995-2004	NS	7/10	[23, 27]
HoBi-like	4 (1.1)	2 (0.8)	2007-2011	S	4/4	[18, 19, 40, 41]

* Herds with different BVDV genotypes or subtypes are counted for each virus type.

** N = Northern Italy, C = Central Italy, S= Southern Italy, I = Islands.

The North macroarea, where cattle population density is highest, accounted for 298 (80.3%) BVDV sequences, the Center for 27 (7.3%), the South for 15 (4%), and the Islands (Sicily and Sardinia) for 31 (8.4%). When differentiated by geographic distribution, BVDV-1 and BVDV-2 were present in the North, BVDV-1 in the Center and the Islands, and BVDV-1, BVDV-2 and HoBi-like virus in the South. Nine different BVDV-1 subtypes were detected in the North, six in the Center and the South, and four in the Islands. The more frequent subtypes (BVDV-1b and 1e) were distributed across the entire country, whereas the less frequent subtypes (BVDV-1a, d, g, h, k) were present in two or three macroareas, except for BVDV-1f, which was limited to the North (Table 1 and Figure 2). The most frequent subtypes were also distributed across all the years, whereas the lower prevalence subtypes and genotypes were detected sporadically (Supplementary Table 1, see Annex section).

Figure 2. Frequency of BVDV genotypes and subtypes in the four Italian macroareas.



The percentage of sequence similarity of pairwise evolutionary distances within BVDV genotypes and subtypes ranged from 83.1% to 100%. Forty of the 49 herds with more than one sequence had the same genotype or subtype (37 BVDV-1, two BVDV-2, one HoBi-like). In addition, 100% sequence identity of BVDV-1within herds was observed in 19/37 herds within a range of sampling between 1 day and 14 months. In 18/37 herds, the nucleotide percentage similarity was ≤ 99.4% within a range of sampling between 2 months and 12 years (Table 2). The remaining nine herds had different genotypes or subtypes, without showing any relationship between the number of sequences analyzed or the time of sampling (Table 3).

Table 2. Sequence similarity (%) of pairwise distances within herds with ≥2 animals/sequences belonging to the same BVDV-1 subtypes.

Sequence similarity (%)	BVDV-1 subtype	Herd identification	Sequences in each herd (n)	Temporal range of collection (months)	Herd production system*
100	1b	RG7	2	2	nd
100	1b	BG23	2	12[a]	D
100	1b	PD3	2	12[a]	nd
100	1b	CO2	2	1[b]	D
100	1b	CN29	2	1[b]	nd
100	1b	AT3	2	1[b]	B
100	1b	CN8	2	1	B
100	1b	TO3	2	1[b]	D
100	1b	TO17	3	13	M
100	1d	CN20	2	1[b]	B
100	1e	RG2	2	8	nd
100	1e	CR22	2	1[b]	D
100	1e	AL2	2	1[b]	M
100	1e	CN24	2	6	D
100	1e	TO8	5	1[b]	D
100	1e	TO14	8	12	B
100	1e	TO19	7	1[b]	D
100	1f	CN2	3	14	M
100	1h	NO1	2	1	D
99.4-100	1e	VI7	4	12[a]	nd
99.4	1b	RG6	2	1[b]	nd
99.4	1b	TO23	2	3	D
99.4	1d	CN19	2	1[b]	D
99.4	1f	AT2	2	1[b]	B
98.8-100	1b	LO2	4	3	D
98.8-100	1f	CN1	7	14	D
98.8	1f	CN5	2	2	D
98.2-100	1b	RG4	4	2	nd
98.2	1b	RM1	2	12[a]	D
98.2	1b	TR1	2	12[a]	B
97.1	1b	CR13	2	12[a]	D
96.5	1b	RM2	2	12[a]	D
94.7-100	1b	LC1	5	129	D
93.6-100	1e	CN16	3	10	B
93-100	1b	RG3	3	3	nd
90.1	1e	CN17	2	13	D
88.9-100	1e	TO12	7	19	M

* D= dairy herd, B= beef herd, M= mixed farm, nd=not determined

[a] data available only for year of sample collection, [b] samples collected on the same day

Table 3. Herds with ≥2 animals/sequences belonging to different genotypes and subtypes.

Genotype	Herd identification	Total sequences No.	Temporal range of collection (months)	Herd production system *
BVDV-1b, 1e	TO9	2	1**	D
BVDV-1b, 1e	CN7	20	1	B
BVDV-1b, 1f	CN6	2	3	B
BVDV-1b, 1k	CS2	3	12	D
BVDV-1b, 2	BS33	2	1**	D
BVDV-1d, 1e	NO2	3	21	M
BVDV-1d, 1e	MN4	2	1**	D
BVDV-1e, 1f	CN3	5	1**	D
BVDV-1e, 1h	LC2	2	19	D

* D= dairy herd, B= beef herd, M= mixed farm
** samples collected on the same day

DISCUSSION

Molecular typing studies have demonstrated a wide genetic diversity of BVDV in Italy [23, 24, 25, 26, 27, 28], where its geographic distribution is influenced by animal movement within the country and/or importation from other countries [21]. Besides the two known BVDV species, a third species, referred to as HoBi-like or BVDV-3, has recently been detected in two herds in Italy [18, 19]. Surveillance of Italian BVDV isolates is therefore advisable to update data on the national genetic variability for mounting local and European-wide alerts.

A comprehensive collection of new BVDV sequences from all cattle breeding areas in Italy was investigated by phylogenetic analysis and compared against a selection of reference sequences of known genotypes and subtypes retrieved from public genetic databases. A total of 371 sequences have been included over a time span of 18 years (1995-2013) and were analyzed using the neighbor-joining and the maximum

likelihood methods. At the genotype level, BVDV-1 had the highest frequency, followed by sporadic BVDV-2 in a limited area of the North (Lombardy and Emilia Romagna), and only once in the South. In the present study, the most recent detection of BVDV-2 dates back to 2004, encompassing nearly a decade during which the potential risk of transmission by BVDV-2 contaminated biological products had been reported [33, 34]. No novel HoBi-like virus was identified using panpestivirus primers described by Letellier *et al.* [29] that are able to detect HoBi-like virus [35; 36]; moreover in central and southern Italy, a protocol for atypical pestivirus detection [14] has been also applied. Regarding Sicily, we cannot exclude failure of HoBi-like viruses detection especially in samples with a low viral load, as previously reported for the primer pair used [4; 35]. The sporadic frequency of the HoBi-like virus in Italy was also confirmed by the National Reference Center for pestivirus that did not detect any other HoBi-like strains on testing an additional collection of 450 BVDV field samples (Giammarioli personal communication, 2014).

The genetic diversity of BVDV-1 in Italy is increasing, as compared to previously published findings [26, 27]. Here we report the circulation of three additional subtypes, namely BVDV-1c, 1k, and 1l, the last classified as the French 1l subtype described by [6], accounting for 11 out of the 17 BVDV-1 subtypes recognized worldwide [9, 10]. The increased phylogenetic diversity of BVDV-1 and the presence of new viral subtypes during the years was also reported in other European countries, especially due to the analysis of broader collection of BVDV isolates and/or introduction and movement of cattle from Europe [6; 22]. Different BVDV-1 subtypes are predominant in European countries, concerning the Italian neighboring countries

the most prevalent subtypes are BVDV-1e in France, BVDV-1e and 1h in Switzerland, BVDV-1h in Austria and BVDV-1d and 1f in Slovenia [6; 20; 37; 38]. Interestingly, BVDV-1e is predominant in Piedmont, the Italian region close to the French border that is also characterized by the major commercial introduction of cattle from France.

Four frequency and distribution patterns of BVDV-1 subtypes were identified in Italy: high prevalent subtypes with a wide temporal-spatial distribution (BVDV-1b and 1e); low prevalent subtypes with a widespread geographic distribution (BVDV-1a, 1d, 1g, 1h, 1k); low prevalent subtypes in restricted geographic areas (BVDV-1f in the North); and sporadic subtypes detected only in single herds (BVDV-1c, 1j and 1l). The North macroarea showed the highest genetic variability, with nine out of 11 BVDV-1 subtypes and the co-circulation of BVDV-2, confirming the predominant role of this area in BVDV introduction into Italy from other European countries [21]. Nevertheless, the identification of some sporadic genetic variants restricted to the Center (BVDV-1l) or the South (BVDV-1c and HoBi-like) and the presence of eight BVDV genotypes and subtypes in the South, despite the low frequency of total sequences, suggest that BVDV has been likely introduced in Italy also through different commercial livestock flows or the use of contaminated biological products.

The genetic variability among BVDV isolates of the same subtype in the same herd was investigated in all the herds where more than one sequence was available. Two major genetic patterns were observed: the presence of herd-specific strains, also for prolonged periods (up to 14 months) and the presence of genetic variability of the same BVDV subtype within single herds, particularly within several months after the first date of sampling, likely indicating a new infection with a different strain. In this

respect, further molecular analysis and investigation of epidemiological links among farms are needed to assess and gain insight into the frequency of re-infection and/or the molecular evolution of BVDV strains detected in the same herd and between herds, as recently applied by [39]. Moreover, a third less frequently observed genetic pattern was the presence of different subtypes or genotypes within the same herd at the same date of sampling, indicating BVDV co- circulation, possibly through exposure to multiple viral sources. Co-circulation of different BVDV subtypes was detected in both milk and beef production systems, confirming that the diversity of viral strains in the Italian cattle population influences the variability also at the herd level.

CONCLUSION

This comprehensive overview of the genetic variability of BVDV strains circulating in Italy highlights a marked genetic diversity. The temporal-spatial distribution of BVDV variants suggests the risk of different routes of BVDV introduction and dispersion, through different commercial livestock flows or the use of contaminated biological products, likely related to the lack of coordinated control measures. Also, it highlights the importance of phylogenetic studies for genetic characterization and for reconstruction of the evolutionary relationships between strains through which the ecological and epidemiological mechanisms driving such genetic heterogeneity may be elucidated. Advanced phylogenetic analysis of the evolutionary dynamics of BVDV strains present in a population can aid in tracing transmission chains, prevent and control infections and sources of re-infections.

CHAPTER 2

CRIMEAN-CONGO HEMORRHAGIC FEVER

BAYESIAN PHYLOGEOGRAPHY OF

CRIMEAN-CONGO HEMORRHAGIC FEVER VIRUS IN EUROPE

PLoS ONE 2013: 8(11): e79663

ABSTRACT

Crimean-Congo hemorrhagic fever (CCHF) is a zoonosis mainly transmitted by ticks that causes severe hemorrhagic fever and has a mortality rate of 5-60%. The first outbreak of CCHF occurred in the Crimean peninsula in 1944-45 and it has recently emerged in the Balkans and eastern Mediterranean. In order to reconstruct the origin and pathway of the worldwide dispersion of the virus at global and regional (eastern European) level, we investigated the phylogeography of the infection by analysing 121 publicly available CCHFV S gene sequences including two recently characterised Albanian isolates. The spatial and temporal phylogeny was reconstructed using a Bayesian Markov chain Monte Carlo approach, which estimated a mean evolutionary rate of 2.96×10^{-4} (95%HPD=1.6 and 4.7×10^{-4}) substitutions/site/year for the analysed fragment. All of the isolates segregated into seven highly significant clades that correspond to the known geographical clades: in particular the two new isolates from northern Albania clustered significantly within the Europe 1 clade. Our phylogeographical reconstruction suggests that the global CCHFV clades originated about one thousand years ago from a common ancestor probably located in Africa. The virus then spread to Asia in the XV century and entered Europe on at least two occasions: the first in the early 1800s, when a still circulating but less or non-pathogenic virus emerged in Greece and Turkey, and the second in the early 1900s,

when a pathogenic CCHFV strain began to spread in eastern Europe. The most probable location for the origin of this European clade 1 was Russia, but Turkey played a central role in spreading the virus throughout Europe. Given the close proximity of the infected areas, our data suggest that the movement of wild and domestic ungulates from endemic areas was probably the main cause of the dissemination of the virus in eastern Europe.

INTRODUCTION

Crimean-Congo hemorrhagic fever virus (CCHFV) belongs to the family *Bunyaviridae*, genus *Nairovirus*. It is an enveloped virus with a negative-sense single stranded RNA genome consisting of one small (S), one medium (M) and one large segment (L) that respectively encode the viral nucleocapsid (N), the membrane glycoprotein precursor (GPC), and RNA-dependent RNA polymerase (L) proteins [1,2].

Crimean-Congo hemorrhagic fever is a mainly tick-borne zoonosis that causes sporadic cases and severe outbreaks of acute human disease with a mortality rate of 5-60% depending on the country [3,4]. The genus *Hyalomma*, particularly the *H. marginatum* of *Ixodes* ticks, is the principal vector of CCHFV, and its geographical distribution correlates with the occurrence of CCHF [4]. In addition to *Hyalomma* spp., the transmission of CCHFV has been associated with *Rhipicephalus* (Europe, South Russia), *Boophilus* (Bulgaria, Russia, Pakistan), *Dermacentor* (Europe), *Ixodes* spp. (Crimea, Moldavia, Bulgaria, Hungary) [5], *Argas persicus* (Uzbekistan) and *Ornithodoros lahorensis* (Iran) [6,7]. Various wild and domestic mammals act as hosts for feeding

ticks that can be infected with the virus; the animal infection is generally asymptomatic.

The disease's seasonal pattern (between spring and early autumn) reflects the high degree of tick activity during this period [8]. Other modes of transmission are direct contact with infected animal blood, viremic CCHF patients and nosocomial infection [9,10].

CCHFV is endemic in parts of Eurasia and Africa, and continues to spread. A number of publications have demonstrated the presence of the virus in about 18 countries in Africa, 11 in South Asia and the Middle East, and eight in eastern Europe [6,8,11].

The migration or transportation of tick-infested birds [4], movements of livestock [12], the environment, tick density, the number of host vertebrate animals [13], and climatic and agricultural changes have all played major roles in taking CCHFV into new areas [8]. The first outbreak of severe hemorrhagic fever associated with CCHFV in Europe (and the world) occurred in the Crimean peninsula (on the Ukrainian coast of the Black Sea) in 1944-45, when hundreds of Soviet soldiers were infected [14]. Other large outbreaks have subsequently been described in Bulgaria (since 1953) and in the regions of Astrakhan and Rostov in south-eastern European Russia [8]. Recent studies in Greece have shown that the seroprevalence of already circulating and new-entry forms has increased from 1.1% to 3.14% over the last 20 years [15]; most of the detected antibodies were against the AP92 strain (isolated in 1970s), which is genetically different from all of the other strains and seems to be less or non-pathogenic to humans. Since 2002, Turkey has reported an exponentially growing number of human cases, whereas seroprevalence has decreased in Bulgaria since the

1974 introduction of an immunisation programme targeting military personnel and healthcare workers [8]. Kosovo experienced a reactivation of natural foci and the re-emergence of CCHF during the 1990s because of dramatic changes in social, political, environmental and economic factors that may together have led to more suitable ecological conditions. There is a seroprevalence of about 24% in the general population living in endemic areas in Kosovo [8,16]. In 2001, there was an outbreak of eight confirmed CCHF cases in Albania at the same time as an epidemic in Kosovo. During 2002-2006, another 24 patients were found to have confirmed CCHF, many of whom lived in Kukës Prefecture, in north-eastern Albania, which shares a border with Kosova. Another outbreak in the same prefecture was reported in 2010, and a few cases were also detected in the south-eastern part of the country in Kolonja near the border with Greece. Recent epidemiological surveillance has found an overall national seroprevalence in Albania of 2.1%, with significant differences between districts ranging from 0.8 to 12% (Silvia Bino, personal communication, 2013).

Analyses of the viral S and L-RNA segments have led to the identification of seven major genetic clades that segregate on a geographical basis [1,17-22]: Africa 1 (also classified as genotype 7), Africa 2 (genotype 5), Africa 3 (genotype 3), Asia 1 (genotype 1), Asia 2 (genotype 2), Europe 1 (genotype 4) and Europe 2 (genotype 6). Analysis of the M segment revealed only six clades, with all of the Asian isolates in a single monophyletic group [23].

We phylogenetically analysed publicly available S region sequences of CCHFV isolated in different geographical areas at different times (including two recently characterized isolates obtained in Albania) in order to reconstruct the most probable places of origin

and pathways of dispersion of the CCHFV clades, particularly concentrating on the strain causing the recent outbreaks in eastern Europe.

MATERIALS AND METHODS

Sequence data sets

The analysis included a total of 121 S gene sequences of CCHFV isolated in various countries of the world and retrieved from public databases (Genbank at http://www.ncbi.nlm.nih.gov/genbank/). The sampling dates ranged from 1958 to 2010, and the sampling locations were Albania (n=2), Turkey (n=37), Russia (n=5), China (n=15), Central Asia (n=4), Iran (n=14), Iraq (n=1), Oman (n=1), Africa (n=31), Pakistan (n=3), Afghanistan (n=1), Kosovo (n=3), Bulgaria (n=2) and Greece (n=2). Accession numbers and main characteristics of the isolates are reported in Table S1. The two Albanian sequences isolated during the recent outbreaks, have been characterized and submitted to GenBank (assigned accession numbers: KC846093 for AL6@03 and KC846094 for AL5@04).

All of the sequences were aligned using ClustalX software [24] and then manually edited using Bioedit software v. 7.0 (freely available at http://www.mbio.ncsu.edu/bioedit/bioedit.html).

Likelihood mapping

In order to obtain an overall impression of the phylogenetic signal present in the analysed CCHFV S sequences, we made a likelihood-mapping analysis of 10,000 random quartets generated using TreePuzzle [25]. A likelihood map consists of an equilateral triangle: each dot within the triangle represents the likelihood of the three

possible unrooted trees for a set of four sequences (quartets) randomly selected from the dataset. The dots close to the corners and at the sides respectively represent tree-like (fully resolved phylogenies in which one tree is clearly better than the others) and network-like phylogenetic signals (three regions in which it is not possible to decide between two topologies); the central area of the map represents a star-like signal (the region in which the star tree is the optimal tree).

Phylogenetic reconstruction

The evolutionary model that best fitted the data was selected using an information criterion implemented in JmodelTest [26], which is freely available at http://darwin.uvigo.es/software/jmodeltest.html.

The phylogenesis of the S gene sequences was initially analysed using a maximum likelihood approach with a new hill-climbing algorithm implemented in the Phyml server v.3.0 (http://www.atgc-montpellier.fr/phyml/) [27]. The reliability of the observed clades was established on the basis of an internal node bootstrap support value of more than 0.90 (after 200 replicates).

The evolutionary rates were estimated using a Bayesian Markov Chain Monte Carlo (MCMC) method implemented in BEAST 1.7.4 [28] with a strict and relaxed molecular clock and an uncorrelated log normal rate distribution model (assuming the GTR+G+I model of nucleotide substitution). As coalescent priors, we compared three parametric demographic models of population growth (constant size, exponential growth, and logistic growth) and a a piecewise-constant Bayesian skyline plot (BSP) [29].

The chains were conducted until reaching convergence as described below, and sampled every 10,000 steps. Convergence was assessed on the basis of the effective

sampling size (ESS=>200) after a 10% burn-in using Tracer software v. 1.5 (http://tree.bio.ed.ac.uk/software/tracer/). Uncertainty in the estimates was indicated by the 95% highest posterior density (95% HPD) intervals and the best fitting models were selected using a Bayes factor (BF) with marginal likelihoods implemented in BEAST. In accordance with [30], the strength of the evidence against H0 was assessed as 2lnBF <2 no evidence; 2–6 weak evidence; 6–10 strong evidence, and >10 very strong evidence. A negative 2lnBF indicates evidence in favour of H0. Only values of >6 were considered significant. The trees were summarised in a maximum clade credibility (MCC) tree (the tree with the largest product of posterior clade probabilities) after a 10% burn-in using the Tree Annotator program included in the BEAST package. The time of the most recent common ancestor (tMRCA) estimates were expressed as mean and 95% HPD years before the most recent sampling dates (corresponding to 2010 in this study).

Bayesian phylogeography

The geographical analysis was made using the continuous time Markov chain (CTMC) process over discrete sampling locations implemented in BEAST, and the Bayesian Stochastic Search Variable Selection (BSSVS) model, which allows diffusion rates to be zero with a positive prior probability. Comparison of the posterior to prior odds that the individual rates are non-zero provides a formal BF for testing the significance of the linkage between locations [31]. Rates with a BF of > 3 were considered well supported and formed the migration pathway.

The MCC tree was selected from the posterior tree distribution after a 10% burn-in using the TreeAnnotator program, version 1.7.4. The final tree was visualised using

FigTree version 1.4 (available at http://tree.bio.ed.ac.uk/software). The significant migration rates were analysed and visualised using SPREAD [32], which is available at http://www.kuleuven.be/aidslab/phylogeography/SPREAD.html. The 121 isolates were assigned to a total of 11 distinct geographic groups corresponding to: Africa, Albania, central Asia, Bulgaria, China, Greece, Kosovo, the Middle East (grouping the 16 isolates from Oman, Iran and Iraq), Pakistan (including 4 sequences from Pakistan and Afghanistan), Russia and Turkey.

In order to provide a spatial projection, the migration routes indicated by the tree were visualised using Google Earth (http://earth.google.com).

In order to investigate a possible bias due to the uneven sample size at each location we performed a sensitivity test involving the randomisation of the tip-location throughout the MCMC procedure (described in more detail in the Figure S1, see Annex section).

The hypothesis of restricted gene flow among distinct CCHFV populations in eastern Europe was investigated in more detail by analysing the European clade using a modified version of the Slatkin and Maddison test [33], which counts migrations to/from different viral subpopulations, using the MacClade version 4 program (Sinauer Associates, Sunderland, MA). A one-character data matrix is obtained from the original data set by assigning a one-letter code to each taxon in the tree indicating its country (or geographic region) of origin, and then the putative origin of each ancestral sequence (i.e. internal node) in the tree is inferred by finding the most parsimonious reconstruction (MPR) of the ancestral character. The final tree length (i.e. the number of observed migrations in the genealogy) can easily be computed and compared with

the tree-length distribution of 10,000 trees obtained by means of random joining-splitting. Observed genealogies that are significantly shorter than the random trees indicate the presence of subdivided populations with restricted gene flow. Specific migrations from/to different countries (character states) were traced using the state changes and stasis tool (MacClade software), which counts the number of changes in a tree for each pair-wise character state. In the presence of multiple MPRs (as in our data sets), the algorithm calculates the average migration count over all possible MPRs for each pair. The resulting pair-wise migration matrix was then normalised, and underwent a randomisation test using 10,000 matrices obtained from 10,000 random trees (by means of the random joining-splitting of the original tree) in order to assess the statistical significance of the observed migration counts. The S gene sequences of the CCHF sequences were assigned to six distinct European countries: Albania, Bulgaria, Greece, Kosovo, Russia and Turkey.

RESULTS

Likelihood mapping

The phylogenetic noise of the data set was investigated by means of likelihood mapping. The evaluation of 10,000 randomly chosen quartets showed that only 4.3% fell in the central area of the likelihood map, and 91.8% were at the corners of the triangle, which suggested that the alignment contained sufficient phylogenetic information (Figure S2, see Annex section).

Phylogenetic analysis

Maximum likelihood analysis of the S segments showed that the isolates segregated into significant clades (Figure S3, see Annex section) corresponding to the seven known clades: Asia 1 Asia 2, Africa 1-3, and Europe 1 and 2. The clades were significantly supported by bootstrap values of ≥90%. As previously described, the Africa 1 clade included mainly (but not exclusively) isolates from western-Africa (Senegal and Mauritania); the Africa 2 clade isolates from central Africa (Uganda, Democratic Republic of the Congo); and the Africa 3 clade isolates from South Africa. The Asia 1 clade included isolates from the Middle East (Oman, Iraq, Pakistan, Afghanistan and Iran), and the Asia 2 clade isolates from the Far East and central Asia (China, Tajikistan, Uzbekistan). The Europe 1 clade included the majority of the eastern European isolates (Russia, Turkey and a single Greek strain), and contained a monophyletic group consisting of the three sequences from Kosovo and the two new Albanian isolates, whereas the Europe 2 clade isolates encompassed the prototype Greek strain AP92 and five Turkish isolates.

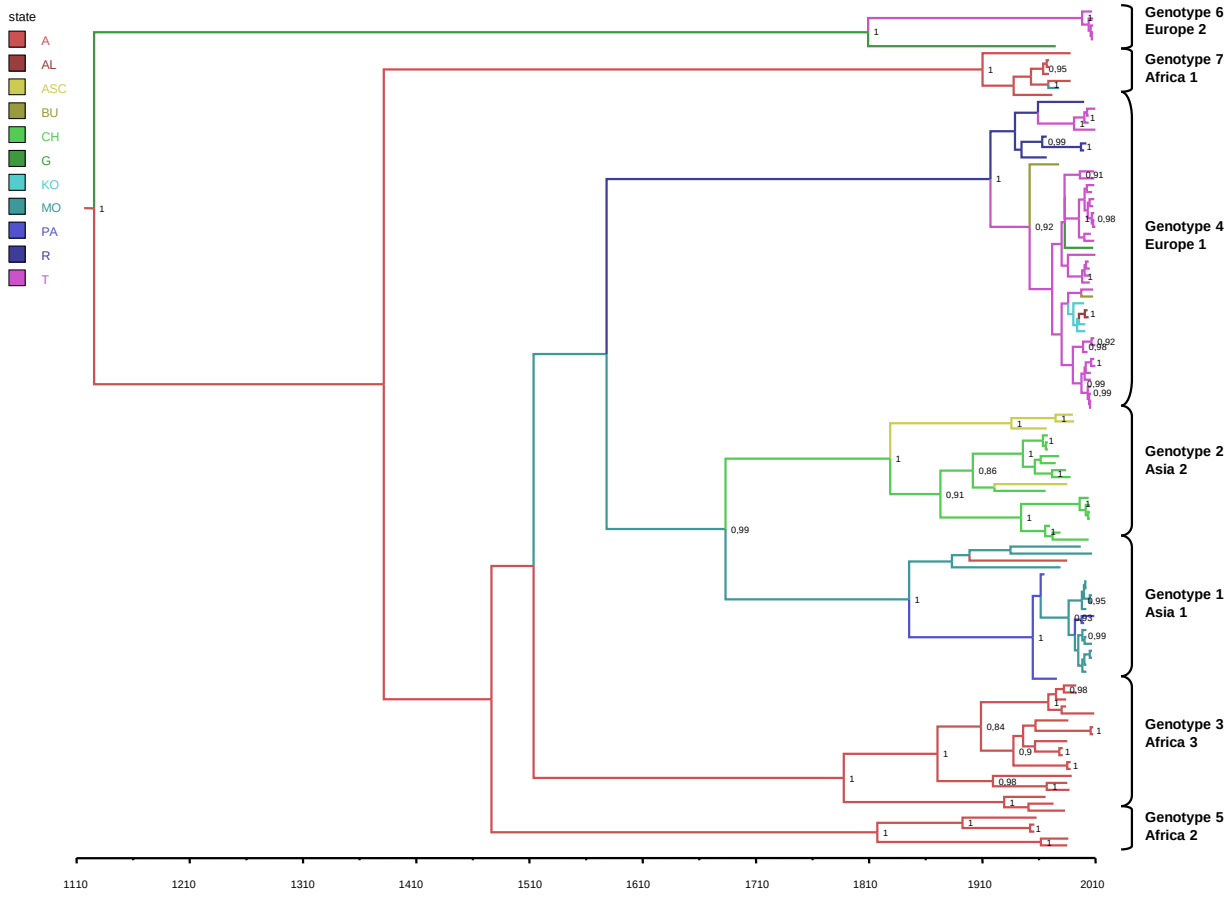
Evolutionary rate estimation and dated tree reconstruction. Comparison of the coalescent models using the BF test showed that BSP (2lnBF constant *vs* BSP= 51.9) under the relaxed clock (2lnBF strict *vs* relaxed clock = 23.5) was the model that best fitted the data. A mean evolutionary rate of 2.96 x 10^{-4} substitutions/site/year (credibility interval 1.6-4.7 x 10^{-4} substitutions/site/year) was estimated under these conditions.

Figure 1 shows the maximum clade credibility (MCC) tree summarising all of the trees obtained during the MCMC search. All of the seven main clades corresponding to the

viral genotypes were confirmed in the Bayesian tree, and supported by posterior probabilities of 1. There were also some highly significant subclades: two main groups in the Europe 1 clade separated the Russian isolates from most of the Turkish and all of the Balkan sequences (the last indicated by blue branches in Figure 1); and subclades in the Asia 2 clade separated the Chinese from the central Asian isolates. Asia 1 and 2 shared a common internal node with a high level of significance in the Bayesian tree (pp=0.99), but not in the ML tree (bootstrap value 0.59).

Using the above evolutionary rate estimates, we calculated the mean tMRCA for the root and each internal node of the tree (Table 1). The mean tMRCA estimate for the tree root was 884.6 years ago (95%HPD 345-1522 years ago), thus suggesting that CCHFV has a remote origin. The Asian clades originated in the first decades of the 1800s, and a common ancestor of the two clades existed a mean 327 years ago (1683). The tMRCA estimates for the African clades were between 100 and 222 years ago, which suggests that these genotypes radiated between the late 1700s and the early 1900s. Finally, the Europe 1 clade had a mean tMRCA of 93 years ago, thus indicating its relatively recent emergence (mean estimate 1917), and the Europe 2 clade had a mean tMRCA of 201 years ago, corresponding to the first decade of the 1800s.

Figure 1. The maximum clade credibility (MCC) tree of CCHFV S gene sequences. The branches are coloured on the basis of the most probable location of the descendent nodes (A=Africa, AL=Albania, ASC=Central Asia, BU=Bulgaria, CH=China, G=Greece, KO=Kosovo, MO=Middle East, PA=Pakistan, T=Turkey). The numbers on the internal nodes indicate significant posterior probabilities (pp>0.8), and the scale at the bottom of the tree represents the number of years before the last sampling time (2010). The main geographical clades (genotypes) have been highlighted.



Phylogeographical analysis

As shown in Figure 1, the highest posterior probability for the location of the tree root was Africa (pp=0.4 *vs* pp= 0.19 for the Middle East, the second most probable location).

The most probable locations for the different clades are shown in Table 1. In particular, the MRCA shared by the two Asian clades was located in the Middle East (pp=0.36 *vs* 0.25 for Pakistan), as was that of the clade Asia 1 (pp=0.57 vs. pp=0.34 for Pakistan). On the contrary, the MRCA of the second Asian clade (Asia 2) was located in China (pp=0.55 *vs* pp=0.38 for central Asia). Interestingly, the MRCA of the Europe 1 clade was more probably located in Russia (pp=0.53 *vs* pp=0.22 for Turkey). The majority of the Turkish isolates segregated into a specific subclade that also included the isolates from Kosovo and Albania (pp=0.55). A Bulgarian isolate was at the outgroup of the Turkish MRCA. Finally, the Europe 2 clade most probably originated in Greece (pp=0.35 *vs* 0.23 for Turkey).

Table 1 Time of the most recent common ancestor (tMRCA) estimates.

CLADE	Subclade	tMRCA[1]	CI tMRCA L[2]	CI tMRCA U[3]	Locality	State pp[4]
ROOT		884,6	345	1522	AFRICA	0,37
AFRICA 1		164,8	94	302	AFRICA	0,98
AFRICA 2		100	51	166	AFRICA	0,99
AFRICA 3		222,5	109	352	AFRICA	0,98
ASIA 1		164,8	79	261	MIDDLE EAST	0,57
	PAKISTAN	55,5	45	72	PAKISTAN	0,9
	MIDDLE EAST	127	56	204	MIDDLE EAST	0,68
ASIA 2		181,5	97	272	CHINA	0,55
	CHINA	137,29	81	205	CHINA	0,85
	CENTRAL ASIA	74,4	46	109	CENTRAL ASIA	0,93
EUROPE 1		93	57	138	RUSSIA	0,53
	TURKEY	58,5	36	86	TURKEY	0,5
	RUSSIA	71,3	49	99	RUSSIA	0,82
EUROPE 2		201	66	373	GREECE	0,35
ASIA 1+2		326,9	168	514	MIDDLE EAST	0,36

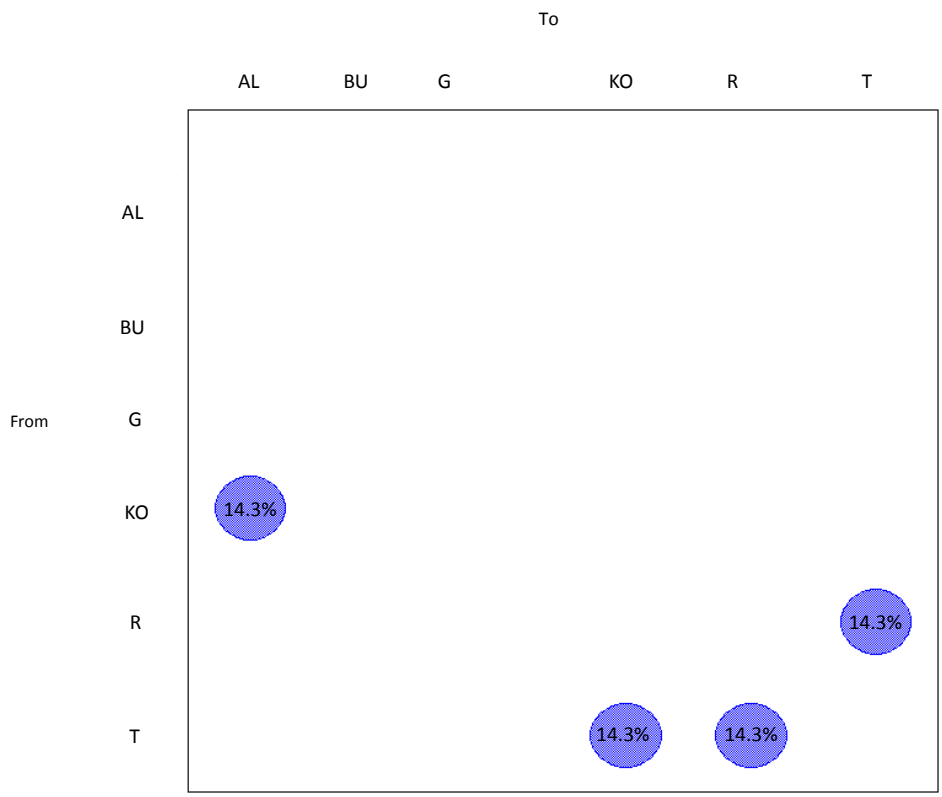
[1]tMRCA: Time of the most recent common ancestor
[2]CI tMRCA L: Credibility Interval Lower
[3] CI tMRCA U: Credibility Interval Upper
[4]pp: posterior probability
Time of the most recent common ancestor (tMRCA) estimates with credibility intervals (95%HPD) with corresponding years and most probable locations with state posterior probabilities (pp) of the main clades observed in the MCC tree of Figure S2.

Bayesian phylogeographical analysis found a mean of 10.5 non-zero rates (95%HPD=10-12), and the rates with a BF of >3 were those between Africa and the Middle East (BF=48.9), the Middle East and Pakistan (BF=2518), China and Central Asia (BF=91), Russia and Turkey (BF=4.7), Turkey and Kosovo (BF=6.0), Kosovo and Albania (BF=12.1), and Bulgaria and Turkey (BF=16.4). The migration routes are shown in Figure 2.

Figure 2. Significant non-zero CCHFV migration rates worldwide.
Rates supported by a BF of >3 are highlighted: the relative strength of the support is indicated by the colour of the lines (from dark red = weak to light red = strong). Dotted lines indicate non-significant linkages. The map was reconstructed using SPREAD (see Methods). The numbers indicate the mean estimated year in which the virus entered the area.

The gene flow (migration) of CCHFV between different European areas was also investigated by analysing the European clade alone using a modified version of the Slatkin and Maddison method. The null hypothesis of panmixia (i.e. no population subdivision or the complete intermixing of sequences from different geographical areas) was tested using a bubblegram (Figure 3), and was rejected for Kosovo, Russia, Turkey and Bulgaria by the randomisation test ($p = 0.0001$). Almost 14.3% of gene flow was from Kosovo to Albania and from Turkey to Kosovo. Interestingly, there was a bidirectional gene flow between Turkey and Russia.

Figure 3 Phylogeographic mapping of CCHF S gene sequences . The bubblegrams show the frequency of gene flows (migrations) to/from ten European countries (same code as that used in Figure 1) . The surface of each circle is proportional to the percentage of observed migrations in the ML genealogy. The migrations were inferred using a modified version of the Slatkin and Maddison algorithm.



DISCUSSION

Various authors have studied the genetic heterogeneity of CCHFV worldwide and found that the existing isolates can be classified into seven main geographical groups: three African, two Asian and two European [21]. As also shown by us, the three African clades include isolates mainly obtained in western (Africa 1), central (Africa 2) and southern Africa (Africa 3); the two Asian clades encompass isolates from the Middle East and Pakistan (Asia 1) or from China and central Asia (Asia 2). Most of the isolates causing recent outbreaks in eastern Europe are in the Europe 1 clade, whereas Europe 2 has largely divergent strains isolated from ticks in Greece (including the prototype strain AP92) [34] and Turkey [35] that are currently associated with sub-clinical or mild cases. The strong spatial structure of the genetic variability of CCHFV suggested it would be worth studying the migration of the clades phylogeographically in order to reconstruct the possible origin and dispersion pathways of the different viral strains. A study published two years ago [22] used a phylogeographical approach based on maximum parsimony, but new Bayesian methods have recently been developed that allow simultaneous estimates of evolutionary rates and migration routes [31], which are useful for reconstructing the past and recent epidemiological history of highly variable viruses [36-38]

In order to calculate the tMRCAs for the root and internal nodes of the phylogeographical tree using a relaxed molecular clock model, we estimated a mean evolutionary rate of 2.96×10^{-4} for the viral genomic S segment. This suggests that, like the majority of other RNA viruses, CCHFV is characterised by a rapid evolution, at least at the level of the S fragment.

Our phylodynamic and phylogeographical reconstruction suggests that the known CCHFV clades shared a common ancestor that existed about one thousand years ago, most probably in Africa. A possible origin in Western Africa was suggested by our preliminary analyses [data not shown], but the extensive geographical intermixing of the three African clades and the relatively small number of African isolates available prevented us from reaching any more precise conclusions.

In our reconstruction, CCHFV left Africa in the second half of the XVII century, reached the Middle East, and then dispersed in two directions in the early XIX century to form the two Asian clades: one spreading in Iran and Pakistan, and the second in China and central Asia (Uzbekistan, Tajikistan). The virus therefore originally spread in an eastward direction to the Middle East and south-east Asia, crossing an area with a constant presence of CCHF susceptible species and the Himalayan mountains as a barrier to natural dispersion. It has recently been speculated that pathogens spread along the Eurasian ruminant route, as in the cases of foot-and-mouth disease [39,40] and Rinderpest [41].

Two highly divergent CCHFV strains entered Europe on at least two occasions: the first in the early 1800s (when a less or non-pathogenic virus confined to the animal/vector population reached Greece and Turkey), and the second in the first decades of the XX century, when a more pathogenic strain caused human outbreaks in eastern Europe until recently. In our phylogeographical reconstruction, the most probable location of the MRCA of this European clade 1 was Russia, which suggests that this was the gateway through which genotype 4 CCHFV entered Europe in the early 1900s. This partially conflicts with previous findings [22] suggesting that Turkey was the origin of

genotype 4, although our analysis of significant migration rates confirmed the central role of Turkey in spreading the virus throughout Europe. Interestingly, the maximum parsimony analysis showed in- and outflows to and from Turkey and Russia, which suggests a continuous exchange between neighboring areas in the region of the Black Sea (possibly justifying the partial discrepancies between studies), and a complex web of viral introductions/transmissions from Turkey to Kosovo, and from Kosovo to Albania.

A recent study [42] has suggested the importance of anthropogenic factors in the dispersion of tick-borne encephalitis virus (TBEV), and provided evidence that its temporal and spatial distribution was related to the first land road into Siberia and the trans-Siberian railway. It is interesting to note that the entire area between the Black and Caspian Seas in the time span estimated for the origin of CCHFV (the first decades of the XX century) was the theatre of a number of wars (i.e. the Crimean war, World War II, and the Civil war) and mass deportations. The virus later spread to Turkey, where it probably circulated in the wildlife population for a long time (as previously postulated by [22]) as the first human cases of hemorrhagic fever were reported in 2002 [8].

In our reconstruction, the virus spread from Turkey to the Balkans, reaching Kosovo in the 1990s and Albania in the last decade. It is possible to hypothesise that the main cause of its dispersion through eastern Europe was the movement of wild and domestic ungulates carrying infected ticks, although outbreaks of CCHFV infection in South Africa have been associated with the passive transportation of infected ticks by birds [9,43], and the same has recently been suggested as introducing the virus to

Spain [13]. Although the region between Black and Caspian Seas is one of the major migratory routes for birds crossing the African-Eurasian flyways (http://www.birdlife.org/), movements from East to West are extremely rare [44].

Turkey has one of the largest ruminant populations in Europe and the Middle East, and witnesses the movement of large and small ruminants for breeding, transhumance (within and across its borders), slaughter and import/export. The main direction of the flow is from neighboring countries such as Russia and Iran [45] and the regions of eastern and central Anatolia to the West. Movements take place throughout the year, but especially before and during religious festivals, when there is an increase in the export of small ruminants [46]. This large-scale movement of livestock over a short period of time has been previously associated with CCHF, Rift Valley fever and *peste-des-petits ruminants* [47]. Moreover, cross-border transhumance occurs on the Russian border [39], and could explain the results of the maximum parsimony analysis of in- and outflows between Turkey and Russia.

The currently used methods of phylogeographical reconstruction are inherently limited by the availability of sample locations and the numbers of isolates at each location. The sensitivity test performed in this study (which suggested that sampling frequencies had little impact on the root location) cannot exclude the influence of unsampled locations. Nevertheless, the analysed data set included all of the sequences with a known sampling location and year that were available in public databases at the time the study began.

In particular, the scarcity of sequences from Bulgaria prevented us from fully clarifying the country's role in disseminating the infection. Bulgarian Thrace and Thrace as a

whole (a geographical area covering south-eastern Bulgaria, north-eastern Greece and the European part of Turkey) is a high-risk zone for the cross-border spread of animal infectious diseases, as has recently been reported in the case of outbreaks of foot-and-mouth disease in Bulgaria and Greece close to the Turkish border [46]. Moreover, the emergence of human CCHF cases in Thrace over a limited period of time and encompassing three different countries -Turkey [48], Bulgaria and Greece [49,50]- confirm the key role of this area in the spread of infectious diseases.

The findings of this study indicate that continuous surveillance of the CCHF epidemic in Turkey and the entire Thracian area may be very important for monitoring and predicting future CCHF outbreaks in the Balkans.

CHAPTER 3

HEPATITIS E VIRUS

PHYLOGEOGRAPHY AND PHYLODYNAMICS OF EUROPEAN

GENOTYPE 3 HEPATITIS E VIRUS

Infection Genetics and Evolution 2014, 25, 138-143

doi:10.1016/j.meegid.2014.04.016

ABSTRACT

Hepatitis E virus is classified into four genotypes that have different geographical and host distributions. The main cause of sporadic autochthonous type E acute hepatitis in developed countries is genotype 3, which has a worldwide distribution and widely infects pigs. The aim of this study was to make hypotheses concerning the origin and global dispersion routes of this genotype by reconstructing the spatial and temporal dynamics of 208 HEV genotype 3 ORF-2 sequences (retrieved from public databases) isolated in different geographical areas.

The evolutionary rates, time of the most recent common ancestors (tMRCAs), epidemic growth and phylogeography of HEV-3 were co-estimated using a MCMC Bayesian method. The maximum clade credibility tree showed the existence of two distinct main clades: clade A, which consists of only European subtypes (HEV-3e and 3f), and clade B, which consists of European subtype 3c and all of the Asian subtypes (3a, 3b and 3d) sharing a common ancestor, which most probably existed in Asia in 1920s. All of the North American isolates belonged to Asian subtype 3a. On the basis of our time-scaled phylogeographical reconstruction, we hypothesise that after originating in the early 1800s in Europe, HEV reached Asia in the first decades of 1900, and then moved to America probably in the 1970s–1980s. Analysis of the skyline plot showed a sharp increase of the number of infections between the 1980s and 2005,

thus suggesting the intervention of new and highly efficient routes of transmission possibly related to changes in the pig industry.

INTRODUCTION

Hepatitis E virus (HEV) is a non-enveloped virus that belongs to the Hepevirus genus of the *Hepeviridae* family. The viral genome, which is represented by a positive-sense single-stranded RNA of about 7.2 kb, contains three partially overlapping open reading frames (ORFs): ORF1, which encodes a non-structural protein with different enzymatic activities (RNA-dependent RNA polymerase, RNA helicase, protease); ORF-2, which encodes the capsid protein, and ORF-3 (overlapping ORF-2), which encodes a viral protein used in virion morphogenesis and release (Kamar et al., 2012; Romano et al., 2013).

HEV has been classified into four main genotypes (1-4) and a number of subtypes (Lu et al., 2006), and other new genotypes have been recently identified in various mammals (Johne et al., 2010; Takahashi et al., 2011; Zhao et al., 2009). The geographical and host-range distribution of HEV genotypes are different: genotypes 1 and 2 infect humans and are widespread in Asia, Africa and Mexico; genotypes 3 and 4, which infect humans but also pigs, boars and deer, give rise to mainly autochthonous HEV infections in non-endemic high-income countries (Dalton et al., 2008; Meng, 2011). Genotype 4 is restricted to China and Japan, and has only recently been identified as the cause of an outbreak in Europe (Bouamra et al., 2014; Garbuglia et al., 2013), whereas HEV genotype 3 has a worldwide distribution and is widespread among pigs in developed countries (Kamar et al., 2012). There is evidence indicating that HEV-

3 infection can be transmitted by undercooked offal or meat, which highlights the zoonotic nature of the infection (Hoofnagle et al., 2012). The prevalence of HEV antibodies is high in developed countries (21% in the US and 7% in Europe, with a maximum of 16% in the area of Toulouse in France) (Kuniholm et al., 2009; Mansuy et al., 2008), thus suggesting a possibly low rate of symptomatic infections associated with genotype 3. In Italy, the prevalence of anti-HEV antibodies is 3-6%, with autochthonous (not travel related) cases mainly due to genotype 3 representing about 16% of the annually reported cases (Romano et al., 2013).

A Bayesian statistical inference framework that allows the simultaneous reconstruction of the temporal and spatial history of an epidemic on the basis of isolates randomly sampled at known times in different places has recently been developed (Drummond et al., 2012), and used to reconstruct the phylodynamic of some highly variable viruses (Ciccozzi et al., 2011; Pybus et al., 2000; Zehender et al., 2012; Zehender et al., 2008).

A number of studies have tried to reconstruct the epidemiological history of HEV at both global (Purdy and Khudyakov, 2010) and regional level (Japan), but partially conflicting results have been obtained particularly concerning on the times of origin and spread of the virus (Nakano et al., 2012a; Nakano et al., 2012b; Tanaka et al., 2006).

In order to make hypotheses about the origin and global dispersion routes of HEV genotype 3, we used a recently developed Bayesian phylogeography framework (Lemey et al., 2009) to reconstruct the spatial and temporal dynamics of 208 HEV genotype 3 ORF-2 sequences isolated in different geographical areas.

MATERIAL AND METHODS

Sequence data set

The analysis included 208 ORF-2 sequences of all of the HEV genotype 3 isolates that were publicly available at the time the study began: 80 from humans and 128 from swine. The sampling dates (available for 75 isolates) ranged from 1993 to 2010, and the sampling locations were Asia (n=120) including China, Japan, Mongolia, Thailand, South Korea, Kyrgyzstan and Taiwan; Europe (n=64) including Italy, France, Spanish, Austria, Germany, Netherlands, England and Sweden; Africa (n=1) and the United States (n=22). The sequences were retrieved from public databases (GenBank,http://www.ncbi.nlm.nih.gov/genbank/) on the basis of the following inclusion criteria: 1) they had to have been published in peer-reviewed journals (except for the new sequences during this study); 2) they had to be representative of all of the main viral subtypes described at the time of the study; 3) there had to be no uncertainty about the genotype/subtype assignment of each sequence, and all were classified as non-recombinant; and 4) the sampling locality and host species had to be known and clearly established in the original publication. The sequences were aligned using ClustalX software (Thompson et al., 1994), manually controlled, and cropped to the same length using Bioedit software (Tom Hall, 2007 – freely available at http://www.mbio.ncsu.edu/bioedit/bioedit.html).

The accession numbers, sampling localities and characteristics of the isolates included in the data set are summarized in Supplementary Table S1.

Likelihood mapping analysis

The phylogenetic signal of each sequence data set was investigated by means of the likelihood-mapping analysis of 10,000 random quartets generated using TreePuzzle. All of the three possible unrooted trees for a set of four sequences (quartets) selected randomly from the data set were reconstructed using the maximum likelihood approach and the selected substitution model. The posterior probabilities of each tree were then plotted on a triangular surface so that the fully resolved trees fell at the corners and the unresolved quartets in the centre of the triangle (star-tree)(Schmidt et al., 2002). Using this strategy, which has been described in detail elsewhere (Strimmer and von Haeseler, 1997), the data are considered unreliable for phylogenetic inference when more than 30% of the dots are in the centre of the triangle.

Bayesian phylogenetics and phylogeography

After alignment, the simplest evolutionary model best fitting the sequence data was selected by the software JmodelTest v.0.0.1(Posada, 2008). The selected model was General Time Reversible (Rodriguez et al., 1990) with a proportion of invariant sites and gamma-distributed rates among sites (GTR+I+G).

The phylogeny, evolutionary rates, times of the most recent common ancestors (tMRCAs), demographic growth and phylogeography were co-estimated in a Bayesian framework using a Markov Chain Monte Carlo (MCMC) method implemented in the Beast package, version 1.74 (Drummond et al., 2012). As coalescent priors, we compared two parametric demographic models of population growth (constant population size, and exponential growth), and a classical piece-wise constant-multiple change process called Bayesian skyline plot (BSP). As sampling dates were available for

only 75 of the sequences included in the study, the substitution rate was estimated on this subset. Given the large credibility interval and the high level of uncertainty of the estimation due to sampling error, we decided to use external calibration to estimate the tMRCAs for each tree's internal node, and adjusted the mean substitution rate to the values estimated on a larger number of dated HEV-3 sequences (Nakano et al., 2012b) using the same genome fragment as that used by us ($1.35x10^{-3}$ subs/site/years, with a credibility interval of $1.0\text{-}1.7x10^{-3}$).

The MCMC analysis was run until convergence with sampling every 1000^{th} generation. Convergence was assessed by estimating the effective sampling size (ESS) after a 10% burn-in, using Tracer software, version 1.5 (http://tree.bio.ed.ac.uk/software/tracer/), and accepting ESS values of 250 or more. The uncertainty of the estimates was indicated by 95% highest posterior density (95% HPD) intervals. The best fitting models were selected using a Bayes factor (BF, based on marginal likelihoods) implemented in Beast (Suchard et al., 2001). In accordance with Kass and Raftery (Kass and Raftery, 1995), the strength of the evidence against H_0 was evaluated as follows: 2lnBF <2 = no evidence; 2-6 = weak evidence; 6-10 strong evidence; and >10 very strong evidence. A negative 2LnBF indicates evidence in favour of H_0. Only values of ≥6 were considered significant.

A recent improvement in the Bayesian framework makes it possible to model spatial diffusion on a time-scaled genealogy as a continuous-time Markov chain (CTMC) (Lemey et al., 2009) over discrete sampling locations. A Bayesian stochastic search variable selection (BSSVS) model that allows the exchange rates in the CTMC to be zero

with some prior probability was used to find a minimal (parsimonious) set of rates explaining diffusion in the phylogeny.

The classic CTMC non-reversible (symmetrical) substitution model for discrete traits does not make it possible to estimate the direction of diffusion rates between locations and, in order to be able to do this, we also used an asymmetrical (non-reversible) substitution model for transition between locations now implemented in Beast (Edwards et al., 2011) and a maximum parsimony-based method implemented in McClade (described below). Comparison of the posterior to prior odds that the individual rates are non-zero provides a formal BF for testing the significance of the linkage between locations. Rates yielding a BF>3 were considered significant (Lemey et al., 2009).

The obtained trees were summarised by Tree Annotator (included in the Beast package) by choosing the tree with the maximum product of posterior probabilities (maximum clade credibility or MCC) after a 10% burn-in. The most probable location at each node was highlighted by labelling the branches with different state colours. The location-annotated MCC tree was converted to a keyhole markup language file (KLM) suitable for viewing with Google Earth (http://earth.google.com/) in order to visualise the diffusion rates over time.

Version 4 of the MacClade program (Sinauer Associates, Sunderland, MA) was used to test viral genes out/in using a modified version of the Slatkin and Maddison test (Slatkin and Maddison, 1989). A one-character data matrix was obtained from the original data set by assigning a one-letter code to each taxon in the tree to indicate its geographical origin. The putative origin of each ancestral sequence (i.e. internal node)

in the tree was then inferred by finding the most parsimonious reconstruction (MPR) of the ancestral character using the ACCTRAN or DELTRAN option. The final tree length (i.e. the number of observed viral gene flow events in the genealogy) can be easily computed and compared to the tree-length distribution of 10,000 trees obtained by random joining-splitting (null distribution). Observed genealogies significantly shorter than random trees indicate the presence of subdivided populations with restricted gene flow. The viral gene flow (migrations) was traced using the 'State changes and stasis tool' (MacClade software), which counts the number of changes in a tree for each pair-wise character state. Gene flow was also calculated for the null distribution in order to assess whether the gene flow events observed in the actual tree were significantly higher (>95%) or lower (<95%) than the values in the null distribution at p=0.05 level.

RESULTS

Likelihood mapping

The likelihood mapping of 10,000 random quartets showed that more than 89% of the randomly chosen quartets were distributed at the corners of the likelihood map and only 4.5% in the central area, thus indicating that the data set contained sufficient phylogenetic information (Supplementary Figure 1S).

Bayesian phylogeny and phylogeography of HEV

The comparison of five coalescent priors under relaxed and strict molecular clock models (see Methods section) showed that the uncorrelated log-normal relaxed clock (2lnBF strict *vs* relaxed clock=17.7) and the Bayesian skyline plot coalescent model

(2lnBF BSP *vs* constant population size=17.8, and *vs* exponential growth=12.45) were significantly better than all of the alternative models.

Under these conditions, a mean evolutionary rate of 1.8 x 10^{-3} subs/site/year with a credibility interval between 0.2 and 3.6 x 10^{-3} subs/site/year was estimated for the subset of 75 HEV ORF-2 sequences with known sampling dates.

Figure 1 shows the phylogeographical maximum clade credibility (MCC) tree of the entire dataset of 208 sequences. The branches of the tree have been coloured on the basis of the most probable location of their descendent nodes.

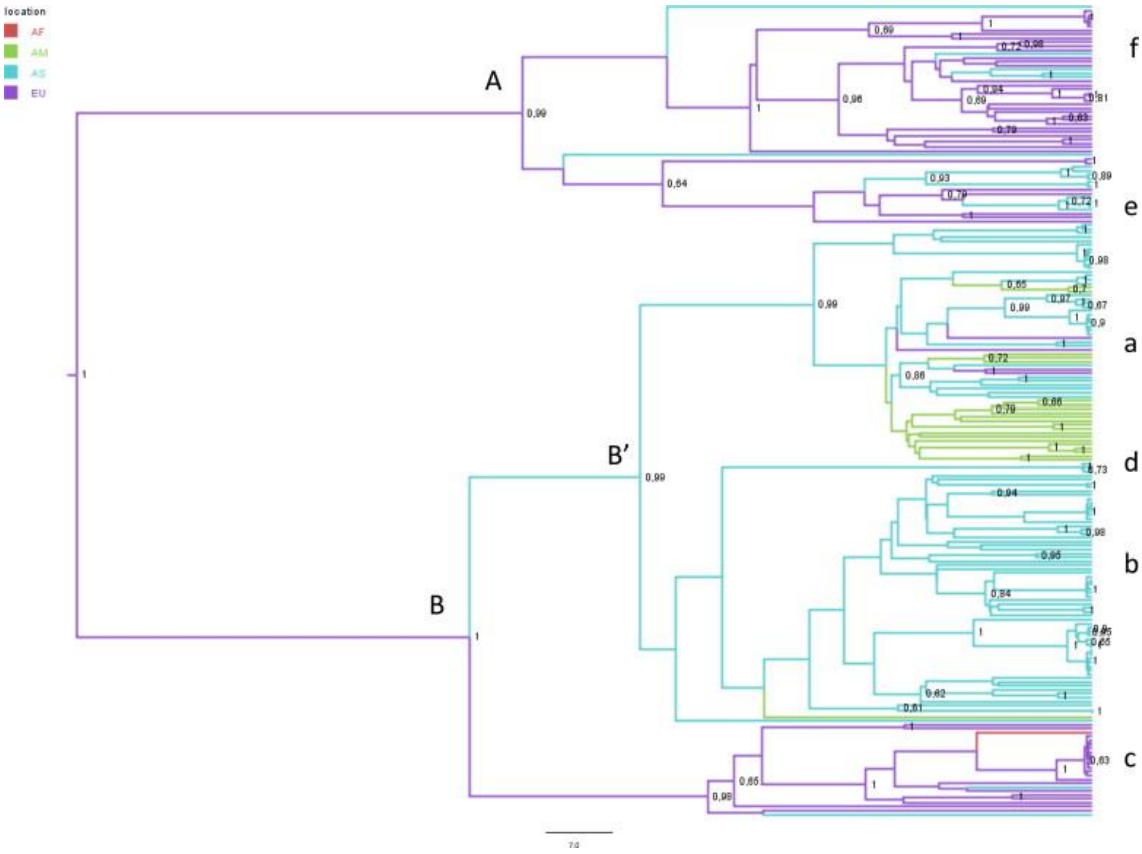
The Bayesian phylogeographical tree showed the existence of two main HEV clades, which we called A (pp=0.98) and B (pp=1). Clade A included two previously described HEV subtypes: HEV-3e (pp=0.64) and HEV-3f (pp=1), whereas clade B included two subclades: one corresponding to subtype 3c (pp=98) and the second (here called B'-pp=0.99) including the three known subtypes HEV-3b, -3a and -3d, although only the last two significantly segregated into distinct monophyletic groups (posterior probabilities of respectively 0.99 and 1.0).

Subtypes 3e and 3f (within clade A) and subtype 3c (within clade B) contained the majority of European isolates, with a number of Japanese sequences mainly included in the subtype 3e. On the contrary, subtypes3a, 3b and 3d included the majority of Asian and all of the North American sequences, which all fell into clade 3a (with the exception of a single Canadian strain classified as 3b) and were largely intermixed with Asian sequences.

The statistical phylogeographical analysis was therefore made by grouping the isolates into four discontinuous geographical areas corresponding to Europe, Asia, North

America and Africa. The analysis showed that the most probable location for the tree root was Europe, with a posterior probability of 0.7 (*vs* the 0.26 posterior probability for Asia); Europe was also the most probable location of the two main clades A and B, as well as subtypes 3e, 3f and 3c. On the contrary, the most probable location of subclade B' and subtypes 3a, 3b and 3d was Asia (with state posterior probabilities of 0.84-1.0) (Fig. 1).

Figure. 1 Bayesian phylogeographical maximum clade credibility (MCC) tree of 208 HEV genotype 3 ORF-2 sequences. The correspondences between the locations and colours are shown in the panel (top left – AF = Africa, AM = America, AS = Asia, EU = Europe); the numbers indicate the posterior probabilities of the corresponding nodes. The main clades and the known HEV-3 subtypes (3a–3f) have been highlighted



Analysis of the dated tree showed that the tMRCA of the tree root (estimated using the external calibration approach, see "Methods" section) was a mean 199.3 years ago (95%HPD 116-289.years ago), which suggests that the currently circulating HEV

genotype 3 shared a common ancestor existing in about 1810. Table 1 shows the estimated tMRCAs of the other significant clades: the two main clades (A and B) originated at the end of 1800s (1889 and 1895, respectively), whereas subtypes 3c and 3e were the earliest genotypes to diverge (about 1917 and 1919), followed by genotype 3f (1935). The common ancestor of all of the Asian subtypes most probably originated in 1920 (B'/subtype 3b), whereas 3a emerged more recently (mean year=1959), and 3d was the last to diverge (in 2002).

Table 1 Posterior probabilities, time of the most recent common ancestor (tMRCA) with credibility intervals (95%HPD), most probable locations with state posterior probabilities (spp) of the main clades observed in the phylogeographic MCC tree (Fig. 1).

Node	pp	Year	Upper 95% HPD	Lower 95%HPD	Location	spp
Root	1.0	1810	1721	1894	Europe	0.70
A	0.99	1889	1838	1939	Europe	0.86
HEV-3f	0.99	1935	1818	1970	Europe	0.99
HEV-3e	0.64	1917	1881	1956	Europe	0.96
B	0.99	1895	1853	1940	Europe	0.60
HEV-3c	0.98	1919	1885	1956	Europe	0.85
B'	0.99	1920	1886	1954	Asia	0.94
HEV-3a	0.99	1959	1944	1980	Asia	0.84
HEV-3d	1.0	2002	2000	2009	Asia	1.00
HEV-3b	NS	1944	–	–	Asia	1.00

Full-size table

Bayesian phylogeography showed a total of 3.7 non-zero rates (95%HPD=3-5), with the linkages between Asia and Europe (BF=5621), Europe and Africa (BF=28.5), and Asia and North America (BF=5621) being statistically well supported. The asymmetrical substitution model for discrete traits, which allows for different rates of migration between locations, showed that the direction was from Europe to Asia (BF 20241) and

Africa (BF=58.9), and from Asia to North America (BF=45.0), whereas there were no significant diffusion rates in the opposite directions (Tab. 2).

Table 2 Significant diffusion rates between localities at Bayesian phylogeographic analysis. Posterior probability of the rate k (pk) and Bayes factor (BF) of the significant migration rates as revealed by reversible (symmetrical) and non-reversible (asymmetrical) substitution models.

Rate (from/to)	**Symmetrical**		**Asymmetrical**	
	pk	**BF**	**pk**	**BF**
Europe/Asia	1.0	5621	1.0	20,241
Asia/America	1.0	5621	0.91	21.8
Europe/Africa	0.98	28.5	0.95	43.44

For the maximum parsimony phylogeographical analysis, the HEV sequences were assigned to four distinct groups labelled in the same way as those used in the Bayesian phylogeographical analysis (Europe, Asia, Africa and North America). The null hypothesis of panmixia (i.e. no population subdivision or complete intermixing of sequences from different geographical areas) was rejected by the randomisation test ($p<0.0001$). The general trend of migration between the four continents was then estimated as observed migration in the tree genealogy (data not shown). HEV gene flow appeared to be asymmetrical, with viral infections expanding from Europe mainly to Asia (2.4%) and much less to Africa (0.2%), whereas the Asian migrations were to Europe (0.5%) and America (1.2%) (Supplementary Figure 2S).

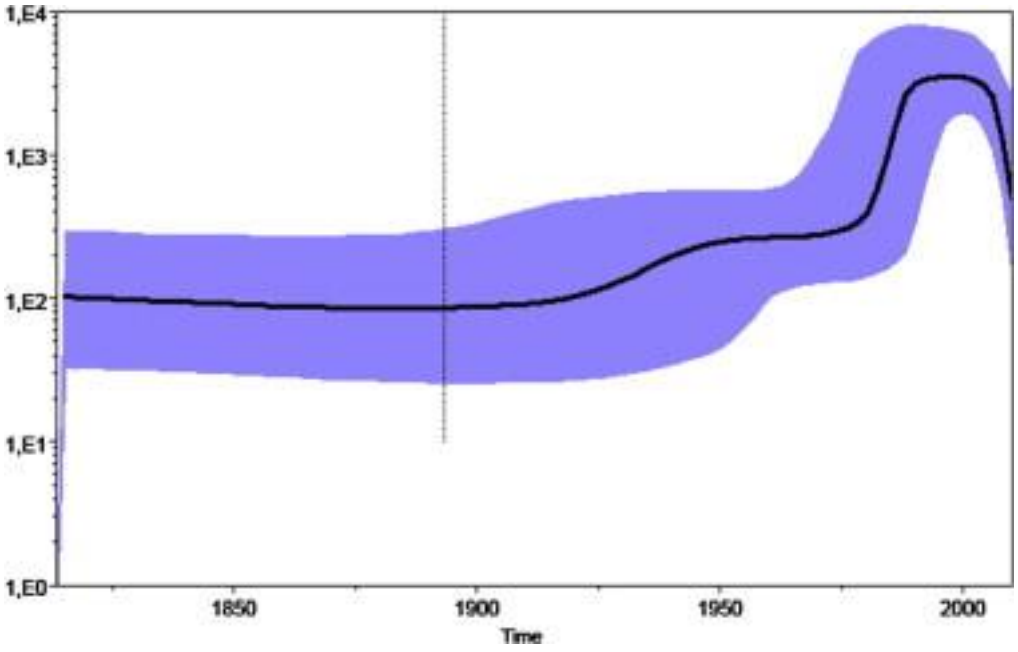
Figure 2 shows the hypothetical worldwide dispersion patterns of HEV-3 as reconstructed on the basis of significant genetic flow rates and dated phylogeny.

Figure 2 Significant non-zero HEV migration rates worldwide. Only the rates supported by a BF of >3 are shown. The map was reconstructed using SPREAD (see Materials and Methods section).

Population dynamics

Analysis of the skyline plot (Fig. 3) showed that the effective number (N_e) of HEV infections at global level remained relatively constant until about the early 1920s and then gradually increased until reaching a first plateau between the 1950s and 1970s. A further sharp increase in the number of infections occurred between the 1980s and 2005, after which the epidemic stopped growing but remained at a level higher than it was at the beginning. The number of infections increased 19-fold from the tree root (1810) to the peak of the curve (early 2000s), and it is now 3.3 times higher.

Figure 3 Bayesian skyline plot of estimates of the effective number of infections (*y* axis; log_{10} scale) over time (*x* axis; calendar years) for HEV isolates, showing the median estimate (solid line) and credibility interval (grey area). The vertical dotted line represents the upper limit of the root height, with the mean tMRCA at the origin.



DISCUSSION

Few previous studies have tried to estimate the speed of evolution of HEV until now. The first was that of Tanaka *et al*. (Tanaka et al., 2006), who used two different internal calibration approaches (based on linear regression analysis and maximum likelihood) and a value of $0.84x10^{-3}$ (0.81-0.94) subs/site/year for the ORF1 RdRp coding fragment; the second was that of Purdy *et al*. (Purdy and Khudyakov, 2010), who used a Bayesian approach similar to that used in our study and obtained a mean evolutionary rate of $1.13x10^{-3}$ for the non-overlapping fragment of ORF2 from a data set of 48 isolates that included several genotypes; and the third was that of Nakano *et al*. (Nakano et al., 2012b), who obtained a mean value of $1.3x10^{-3}$ with a credibility interval of $1.0\text{-}1.7x10^{-}$

^{3} subs/site/year for the ORF2 sequences from a large data set of indigenous Japanese genotype 3 HEV isolates. However, these studies proposed very different tMRCA estimates for the origin of genotype 3. The difficulty in obtaining homogeneous results is probably attributable to sampling bias (because of the small number of available dated sequences), the different viral genes studied, and possible overlaps of the different ORFs.

Our estimate of the evolutionary rate of ORF2 based on a subset of 75 sequences with known sampling dates gave a similar mean value ($1.8x10^{-3}$ subs/site/year) but with a broad credibility interval (between 0.96 and $3.7x10^{-3}$ subs/site/year) and so, given the affinity between our study and that of Nakano *et al*. (Nakano et al., 2012b) in terms of the methods used and the analysed genes and genotypes, we used their mean evolutionary rate and credibility interval when reconstructing the time scale of the evolution of HEV genotype 3.

On the basis of this spatio-temporal reconstruction, we suggest that the currently circulating HEV genotype 3 originated in Europe in the first decade of the 1800s and diverged in the late XIX century into its two main clades: clade A, which gave origin to the European subtypes 3e and 3f in the first half of the XX century; and clade B, which diverged into the European genotype 3c and the common ancestor of the Asian subtypes around the 1920s. The Asian strain was probably closely related to subtype 3b whereas subtypes 3a and 3d differentiated later: the first about the end of the 1950s, and the second in the early 2000s. Interestingly, a number of Japanese isolates were also included in European subtype 3e, thus suggesting that HEV-3 underwent multiple migration events between Europe and Japan. On the basis of the flow rates

reconstructed phylogeographically, HEV-3 migrated from Europe to Asia, but not *vice versa* (or to a lesser extent), as suggested by the maximum parsimony analysis.

But what was the origin of the Asian strains? It is interesting to note that the consumption of pork was prohibited in Japan after the introduction of Buddhism in the sixth and seventh centuries until the ban was lifted in 1872 by the Meiji Emperor (http://www.kikkoman.co.jp/kiifc/foodculture /pdf_09/e_002_008.pdf). There was a significant increase the importation of pork in the first decades of the 1900s, when it was considered patriotic to eat meat and its consumption was actively encouraged by the central government. As pork was much cheaper to produce than beef, its consumption rapidly increased and, given the relatively scarcity of pigs in Japan at the time, many were imported (particularly from the UK) (Inoue et al., 2006; Nakano et al., 2012a; Nakano et al., 2012b) and explain the arrival of an European zoonosis.

Interestingly, clade 3a included all of the isolates from North America, except for one Canadian isolate belonging to clade 3b. The North American strains did not form a single clade within 3a, but were intermixed with Asian isolates and formed only small significant groups. Our genetic flow reconstructions suggest that subtype 3a originated in Asia and was transported to the USA at various times between the 1970s and 1990s, but not in the opposite direction. However, this finding may have been influenced by the large number of Japanese isolates and the relatively small number of US sequences (reported in a single study), and further more specific studies are needed to clarify this point.

Interestingly, the BSP population dynamics analysis showed that the number of infections worldwide remained constant until the 1950s, and then started to increase

gradually probably as a result of the spread of the subtype in Japan and the USA; however, the number of infections increased 19 times during a 20-year period starting in the 1980s, and it is only over the last decade (since 2005) that it has decreased to the level of the 1990s.

These trends are similar to those of pork consumption in developed countries. The increase in production was highest during the period 1961-1981: 7.7-fold in Japan, 1.7-fold in Europe, and 1.4-fold in Northern America (http://www.fao.org/statistics). During the following twenty years (from 1981 to 2001), production in Northern America and Japan increased by respectively 1.2 and 1.4 times, whereas European production increased only slightly. At the same time, there was a shift to fewer but larger farms, associated with structural changes from partially to totally confined housing. Moreover, in order to increase growth and feed efficiency, spray-dried porcine plasma (SDPP) was widely added to the diets of weanling pigs, although it has banned in the European Union since 2000 as a result of the bovine spongiform encephalopathy crisis. Spray-drying significantly reduces the number of viable micro-organisms (Pujols et al., 2011), but there have been no reports of HEV inactivation in porcine blood products. All of these changes increased productivity and lowered prices, thus leading to a greater density of pigs per farm and a larger amount of manure that influenced the spread of HEV-3 infection on the farms and in the general environment. Together with the increase in pork consumption, this may explain the considerable increase in HEV infection among pigs (and consequently humans) in developed countries during the final decades of the 1900s and the beginning of the present century. Further and more specific studies are needed to verify the role of this

or other causes in the spread of HEV-3 infection among pigs in the last decades in order to prevent the further worldwide dispersion of this and other similar food-transmitted infections.

OVERALL CONCLUSIONS

OVERALL CONCLUSIONS

On the whole, the phylogeographic analyses carried out contributed to reconstruct the origin and spatial-temporal dispersion of the viral diseases investigated.

To highlight the main findings achieved the following summary points are reported:

- Pestivirus – BDV: Pyrenean chamois BDV phylogenetic group originated from sheep BDV genotype 4, generating a founder effect due to intra-species spread and spatial dispersion in Pyrenean chamois, with a mean estimation of the most common ancestor falling in 1988. The pathway of dispersion of isolates suggests a complex exchange between neighboring Pyrenean chamois populations, still going on such as Western direction.
- Pestivirus – BVDV: Italy is one of the countries with the highest genetic diversity of BVDV worldwide. Northern Italy ranked first for BVDV introduction, prevalence, and dispersion. Nevertheless, the presence of sporadic variants in other restricted areas suggests also the risk of introduction through different commercial livestock flow or contaminated biological products.
- CCHFV clades originated about one thousand years ago from a common ancestor probably located in Africa. The virus then spread to Asia and entered Europe on at least two occasions: the first in the early 1800s, less or non-pathogenic genetic variant, the second one in the early 1900s, a pathogenic strain, that began to spread in eastern Europe. The most probable location for the origin of this pathogenic clade was Russia, but Turkey played a central role in spreading the virus throughout Europe.

- HEV: The phylogeographical reconstruction of HEV-3 suggests that after originating in the early 1800s in Europe, HEV-3 reached Asia in the first decades of 1900, and then moved to America probably in the 1970s–1980s. HEV-3 population dynamics showed a sharp increase of the number of infections between the 1980s and 2005, thus suggesting the intervention of new and highly efficient routes of transmission possibly related to changes in the pig industry.

Overall, the number of available sequences to include in the datasets represented the main constrain to carry out the above studies. Indeed the sequence datasets have been generated including not only the new sequences obtained in our studies, but also sequences retrieved from public databases. These sequences to be included need the following information: geographic locality where the strain was isolated and sampling date. The currently used methods of phylogeographical reconstruction are inherently limited by the availability of sequences that respect the minimum criteria of inclusion. Moreover spatial and temporal information can range from a macro level, such as Africa and year of isolation, to a micro level, such as the geographic coordinates and day. The overall information availability will allow to define the aim of the investigation and the reliability of the results. For example, in the study conducted on BDV strains from Pyrenean chamois a high quality sequences dataset has been obtained, indeed sequences were characterized by geographic coordinates and month/year. The intensive monitoring of chamois population allowed to collect a large number of strains from different areas in the Pyrenees (Marco et al., 2009). On the contrary, in the study on CCHFV, the analyzed data set included all the sequences with a known

sampling location and year that were available in public databases at the time the study began, nevertheless the scarsity of sequences from countries such as Bulgaria prevented us from fully clarifying the country's role in disseminating the infection.

Concerning the aspect of genome sequences availability, the new powerful methods for genomic characterization, such as next-generation sequencing, will contribute to increase the availability of whole genome sequences and likely to detect and characterize pathogens with or without prior knowledge of their existence, but mainly viruses that have mutualistic or symbiotic relationships with their hosts (Roossinck, 2011).

On the side of quality of genome database, it is essential that besides the increasing number of viral genome sequences that will be available, a systematic collection of epidemiological information need to be acquired to infer the patterns of viral spread within complex transmission networks (Holmes, 2008).

An essential aspect to achieve a comprehensive inference of phylogeographic analyses is a multidisciplinary approach to understand the inextricable interconnection of humans, domestic animals, and wildlife and their social and ecological environment. This approach has been definitely applied in the studies we performed, involving researchers in biology, human and veterinary medicine.

In conclusion, the reliability of the results obtained and the ability to infer viral evolutionary patterns will depend not only on the increasing availability of sequences databases and the development of statistical methods, but also on an effective interdisciplinary approach.

REFERENCES

REFERENCES ARE LISTED BY SECTIONS

- **REEFERENCES OF ABSTRACT, INTRODUCTION, AIMS, OVERALL CONCLUSIONS**

Avise, JC. 1998 The history and purview of phylogeography: a personal reflection. Molecular Ecology, 7, 371–379.

Chua KB, Bellini WJ, Rota PA, Harcourt BH, Tamin A, Lam SK, Ksiazek TG, Rollin PE, Zaki SR, Shieh W, Goldsmith CS, Gubler DJ, Roehrig JT, Eaton B, Gould AR, Olson J, Field H, Daniels P, Ling AE, Peters CJ, Anderson LJ, Mahy BW 2000 Nipah virus: a recently emergent deadly paramyxovirus. Science 288: 1432–1435.

Cleaveland S, Laurenson MK, Taylor LH. 2001 Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philos Trans R Soc Lond B Biol Sci. 356 (1411): 991-9

Cleaveland S, Meslin FX, Breiman R. 2006 Dogs can play useful role as sentinel hosts for disease. Nature 440: 605.

Cutler SJ, Fooks AR, van der Poel WH. 2010 Public health threat of new, reemerging, and neglected zoonoses in the industrialized world. Emerg Infect Dis. 16 (1): 1-7.

Dalton, HR, Bendall R, Ijaz S, Banks M. 2008. Hepatitis E: an emerging infection in developed countries. The Lancet infectious diseases 8, 698-709.

De Benedictis, P, Gallo, T, Iob, A, Coassin, R, Squecco, G, Ferri, G, D'Ancona, F, Marangon, S, Capua, I, Mutinelli, F. 2008. Emergence of fox rabies in northeastern Italy. Euro Surveill. 13, Pii: 19033.

De Benedictis, P, Capua, I, Mutinelli, F, Wernig, JM, Aric, T, Hostnik, P, 2009. Update on fox rabies in Italy and Slovenia. Rabies Bull. Eur. 33: 5–7.

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W. 2002 Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161 (3): 1307-20.

Duffy S, Shackelton LA, Holmes EC. 2008 Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet. 9 (4): 267-76.

Ergonul O. 2006 Crimean-Congo haemorrhagic fever. Lancet Infect Dis 6: 203-214.

Faria NR, Suchard MA, Rambaut A, Lemey P. 2011 Toward a quantitative understanding of viral phylogeography. Curr Opin Virol. 1(5): 423-9.

Fusaro A, Monne I, Salomoni A, Angot A, Trolese M, Ferrè N, Mutinelli F, Holmes EC, Capua I, Lemey P, Cattoli G, De Benedictis P. 2013 The introduction of fox rabies into Italy (2008-2011) was due to two viral genetic groups with distinct phylogeographic patterns. Infect Genet Evol. 7: 202-9.

Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, Holmes EC. 2004 Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303(5656): 327-32.

Holmes EC. 2004 The phylogeography of human viruses. Mol Ecol. 13(4): 745-56.

Holmes EC. 2008 Evolutionary history and phylogeography of human viruses. Annu Rev Microbiol. 62: 307-28.

Hoofnagle JH, Nelson KE, Purcell RH.2012. Hepatitis E. The New England journal of medicine 367: 1237-1244.

Kidd DM, Ritchie MG. 2006 Phylogeographic information systems: putting the geography into phylogeography Journal of Biogeography 33: 1851–1865

Knobel DL, Cleaveland S, Coleman PG, Fèvre EM, Meltzer MI, Miranda ME, Shaw A, Zinsstag J, Meslin FX. 2005 Re-evaluating the burden of rabies in Africa and Asia. Bull World Health Organ 83: 360–368.

Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009 Bayesian phylogeography finds its roots. PLoS Comput Biol 5: e1000520

Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S, Wang LF. 2005 Bats are natural reservoirs of SARS-like coronaviruses. Science 310: 676–679.

Loeffler F, Frosch P. 1897 Berichte der Kommission zur Erforschungder Maul-und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin. Zent Bakt Parasitkde Abt I 23: 371–391

Longdon B, Brockhurst MA, Russell CA, Welch JJ, Jiggins FM. 2014 The evolution and genetics of virus host shifts. PLoS Pathog. 10(11): e1004395.

Marco I, Rosell R, Cabezón O, Mentaberre G, Casas E, Velarde R, Lavín S. 2009 Border disease virus among chamois, Spain. Emerg Infect Dis. 15(3):448-51.

Marco, I, Cabezón, O, Velarde, R, Fernández-Sirera, L, Colom-Cadena, A, Serrano, E, Rosell, R, Casas-Díaz, E, Lavín, S. 2015 The two sides of border disease in Pyrenean chamois (*Rupicapra pyrenaica*): silent persistence and population collapse. Anim. Health Res. Rev. 16: 70-77.

Nouvellet, P, Donnely, C, De Nardi, M, Rhodes, C, De Benedictis, P, Citterio, C, Obber, F, Lorenzetto, M, Dalla Pozza, M, Cauchemez, S, Cattoli, G, 2013. Rabies and canine distemper virus epidemics in the red fox population of northern Italy (2006–2010). PloS ONE 8, e61588.

Pybus OG, Rambaut A. 2009 Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet. 10(8): 540-50

Roossinck MJ. 2011 The good viruses: Viral mutualistic symbioses. Nat Rev Microbiol. 9: 99–108.

Rosenberg R. 2015 Detecting the emergence of novel, zoonotic viruses pathogenic to humans. Cell Mol Life Sci. 72(6): 1115-25.

Talbi C, Holmes EC, De Benedictis P, Faye O, Nakouné E, Gamatié D, Diarra A, Elmamy BO, Sow A, Adjogoua EV, Sangare O, Dundon WG, Capua I, Sall AA, Bourhy H. 2009 Evolutionary history and dynamics of dog rabies virus in western and central Africa. J Gen Virol. 90: 783-91.

Talbi C, Lemey P, Suchard MA, Abdelatif E, Elharrak M, Nourlil J, Faouzi A, Echevarría JE, Vazquez Morón S, Rambaut A, Campiz N, Tatem AJ, Holmes EC, Bourhy H. 2010 Phylodynamics and human-mediated dispersal of a zoonotic virus. PLoS Pathog. 6(10):e1001166.

Webby RJ, Webster RG 2001 Emergence of influenza A viruses. Philos Trans R Soc Lond B Biol Sci 356: 1817–1828.

Woolhouse M., Taylor LH, Haydon DT. 2001 Population biology of multihost pathogens Science 292 (5519): 1109-1112.

Woolhouse ME&, Gowtage-Sequeria S. 2005 Host range and emerging and reemerging pathogens. Emerg Infect Dis. 11(12): 1842-7.

Woolhouse ME, Haydon DT, Antia R. 2005 Emerging pathogens: the epidemiology and evolution of species jumps. Trends Ecol Evol. 20(5): 238-44.

Woolhouse M, Gaunt E 2007 Ecological origins of novel human pathogens. Crit Rev Microbiol 33:231–242.

Zehender G, Ebranati E, Bernini F, Lo Presti A, Rezza G, Delogu M, Galli M, Ciccozzi M. 2011 Phylogeography and epidemiological history of West Nile virus genotype 1a in Europe and the Mediterranean basin. Infect Genet Evol. 11(3): 646-53.

Zumla A, Hui DS, Perlman S. 2015 Middle East respiratory syndrome. Lancet. 386 (9997): 995-1007.

- **REFERENCES OF CHAPTER 1**

SPATIAL AND TEMPORAL PHYLOGENY OF BORDER DISEASE VIRUS IN PYRENEAN CHAMOIS

Arnal M, Fernández-de-Luco D, Riba L, Maley M, Gilray J, Willoughby K, Vilcek S, Nettleton PF. 2004 A novel pestivirus associated with deaths in Pyrenean chamois (Rupicapra pyrenaica pyrenaica). J Gen Virol. 85 (Pt 12):3653-7.

Becher P, Orlich M, Shannon AD, Horner G, König M, Thiel HJ. 1997 Phylogenetic analysis of pestiviruses from domestic and wild ruminants. J Gen Virol. 78 (Pt 6): 1357-66.

Becher P, Avalos Ramirez R, Orlich M, Cedillo Rosales S, König M, Schweizer M, Stalder H, Schirrmeier H, Thiel HJ. 2003 Genetic and antigenic characterization of novel pestivirus genotypes: implications for classification. Virology. 20;311(1): 96-104.

Braun U, Bachofen C, Büchi R, Hässig M, Peterhans E. 2013 Infection of cattle with Border disease virus by sheep on communal alpine pastures. Schweiz Arch Tierheilkd.155 (2): 123-8.

Cabezón O, Rosell R, Velarde R, Mentaberre G, Casas-Díaz E, Lavín S, Marco I. 2010 Border disease virus shedding and detection in naturally infected Pyrenean chamois (Rupicapra pyrenaica). J Vet Diagn Invest. 22(5): 744-7.

Casaubon J, Vogt HR, Stalder H, Hug C, Ryser-Degiorgis MP. 2012 Bovine viral diarrhea virus in free-ranging wild ruminants in Switzerland: low prevalence of infection despite regular interactions with domestic livestock. BMC Vet Res. 8: 204.

Citterio CV, Luzzago C, Sala M, Sironi G, Gatti P, Gaffuri A, Lanfranchi P 2003 Serological study of a population of alpine chamois (Rupicapra r rupicapra) affected by an outbreak of respiratory disease. Vet Rec. 153(19): 592-6.

Cranwell MP, Otter A, Errington J, Hogg RA, Wakeley P, Sandvik T. 2007 Detection of Border disease virus in cattle. Vet Rec. 161(6): 211-2.

Drummond AJ, Rambaut A 2007 BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007 Nov 8;7:214.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005 Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22: 1185-1192.

Drummond AJ, Suchard MA, Xie D, Rambaut A 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7.Molecular Biology and Evolution 29: 1969-1973.

Dubois E, Russo P, Prigent M, Thiéry R 2008 Genetic characterization of ovine pestiviruses isolated in France, between 1985 and 2006. Vet Microbiol. 130(1-2):69-79.

Fernández-Sirera, L., Cabezón, O., Allepuz, A., Rosell, R., Riquelme, C., Serrano, E., Lavín, S., Marco, I. 2012a Two different epidemiological scenarios of border disease in the populations of Pyrenean chamois (Rupicapra p. pyrenaica) after the first disease outbreaks. PLoS ONE 7, e51031.

Fernández-Sirera L, Riba L, Cabezón O, Rosell R, Serrano E, Lavín S, Marco I. 2012b Surveillance of border disease in wild ungulates and an outbreak in Pyrenean chamois (Rupicapra pyrenaica pyrenaica) in Andorra. J Wildl Dis. 48(4): 1021-9.

Finlaison DS, King KR, Frost MJ, Kirkland PD. 2009 Field and laboratory evidence that Bungowannah virus, a recently recognised pestivirus, is the causative agent of the porcine myocarditis syndrome (PMC). Vet Microbiol. 136 (3-4): 259-65.

Gaffuri A, Giacometti M, Tranquillo VM, Magnino S, Cordioli P, Lanfranchi P. 2006 Serosurvey of roe deer, chamois and domestic sheep in the central Italian Alps. J Wildl Dis. 42(3): 685-90.

Giammarioli M, La Rocca SA, Steinbach F, Casciari C, De Mia GM. 2011 Genetic and antigenic typing of border disease virus (BDV) isolates from Italy reveals the existence of a novel BDV group. Vet Microbiol. 147 (3-4): 231-6.

Hamers C, Dehan P, Couvreur B, Letellier C, Kerkhofs P, Pastoret PP. 2001 Diversity among bovine pestiviruses. Vet J. 161(2):112-22.

Hurtado, A, Garcia-Perez A L. Aduriz G, Juste RA 2003. Genetic diversity of ruminant pestiviruses from Spain. Virus Research 92: 67–73.

Hurtado A, Aduriz G, Gómez N, Oporto B, Juste RA, Lavin S, Lopez-Olvera JR, Marco I. 2004 Molecular identification of a new pestivirus associated with increased mortality in the Pyrenean Chamois (Rupicapra pyrenaica pyrenaica) in Spain. J Wildl Dis. 40(4): 796-800.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of American Statistical Association 90, 773-795.

Kirkland PD, Frost MJ, Finlaison DS, King KR, Ridpath JF, Gu X. 2007 Identification of a novel virus in pigs-Bungowannah virus: a possible new species of pestivirus. Virus Res. 129(1):26-34

Krametter R, Nielsen SS, Loitsch A, Froetscher W, Benetka V, Moestl K, Baumgartner W 2004 Pestivirus exposure in free-living and captive deer in Austria. J Wildl Dis. 40(4):791-5.

Oguzoglu TC, Tan MT, Toplu N, Demir AB, Bilge-Dagalp S, Karaoglu T, Ozkul A, Alkan F, Burgu I, Haas L, Greiser-Wilke I. 2009 Border disease virus (BDV) infections of small ruminants in Turkey: a new BDV subgroup? Vet Microbiol. 135(3-4):374-9.

Olde Riekerink RG, Dominici A, Barkema HW, de Smit AJ 2005 Seroprevalence of pestivirus in four species of alpine wild ungulates in the High Valley of Susa, Italy. Vet Microbiol. 108 (3-4): 297-303.

Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5, e1000520.

Marco, I. 2012 Pestivirus of chamois and Border Disease. In: Infectious diseases of wild mammals and birds in Europe. Gavier-Widén, D., Duff, J.P. and Meredith, A. (eds). Wiley-Blackwell, West Sussex, UK. Pp.: 147-152.

Marco I, Lopez-Olvera JR, Rosell R, Vidal E, Hurtado A, Juste R, Pumarola M, Lavin S. 2007 Severe outbreak of disease in the southern chamois (Rupicapra pyrenaica) associated with border disease virus infection. Vet Microbiol. 120(1-2):33-41

Marco I, Rosell R, Cabezón O, Mentaberre G, Casas E, Velarde R, López-Olvera JR, Hurtado A, Lavín S. 2008 Epidemiological study of border disease virus infection in Southern chamois (Rupicapra pyrenaica) after an outbreak of disease in the Pyrenees (NE Spain).Vet Microbiol. 127(1-2): 29-38. Epub 2007 Aug 17.

Marco I, Rosell R, Cabezón O, Mentaberre G, Casas E, Velarde R, Lavín S. 2009 Border disease virus among chamois, Spain. Emerg Infect Dis 15: 448–451.

Marco I, Cabezón O, Rosell R, Fernández-Sirera L, Allepuz A, Lavín S. 2011 Retrospective study of pestivirus infection in Pyrenean chamois (Rupicapra pyrenaica) and other ungulates in the Pyrenees (NE Spain).Vet Microbiol. 149(1-2):17-22.

Marco, I., Cabezón, O., Velarde, R., Fernández-Sirera, L., Colom-Cadena, A., Serrano, E., Rosell, R., Casas-Díaz, E., Lavín, S. 2015. The two sides of border disease in Pyrenean chamois (Rupicapra pyrenaica): silent persistence and population collapse. Anim. Health Res. Rev. 16, 70-77.

Martin C, Duquesne V, Adam G, Belleau E, Gauthier D, Champion JL, Saegerman C, Thiéry R, Dubois E. 2015 Pestiviruses infections at the wild and domestic ruminants interface in the French Southern Alps.Vet Microbiol. 175(2-4):341-8.

Pioz M, Loison A, Gibert P, Dubray D, Menaut P, Le Tallec B, Artois M, Gilot-Fromont E. 2007 Transmission of a pestivirus infection in a population of Pyrenean chamois. Vet Microbiol. 119(1):19-30

Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol Biol Evol 25, 1253-1256.

Rodríguez F, Oliver JL, Marín A, Medina JR 1990. The general stochastic model of nucleotide substitution J Theor Biol 142, 485-501.

Rosell R, Cabezón O, Pujols J, Domingo M, Muñoz I, Núñez JI, Ganges L. 2014 Identification of a porcine pestivirus as a border disease virus from naturally infected pigs in Spain. Vet Rec. 174(1):18.

Schmidt HA, Strimmer K, Vingron M, von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18, 502-504.

Strimmer and von Haeseler 1997 Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc Natl Acad Sci U S A. 94(13):6815-9.

Suchard MA, Weiss RE, Sinsheimer JS 2001 Bayesian selection of continuous-time Markov chain evolutionary models. Mol Biol Evol 18: 1001-1013.

Thabti F, Letellier C, Hammami S, Pépin M, Ribière M, Mesplède A, Kerkhofs P, Russo P. 2005 Detection of a novel border disease virus subgroup in Tunisian sheep. Arch Virol. 150(2):215-29.

Valdazo-González B, Alvarez-Martínez M, Greiser-Wilke I. 2006 Genetic typing and prevalence of Border disease virus (BDV) in small ruminant flocks in Spain. Vet Microbiol. 117(2-4):141-53.

Valdazo-González B, Alvarez-Martínez M, Sandvik T. 2007 Genetic and antigenic typing of border disease virus isolates in sheep from the Iberian Peninsula. Vet J 174(2): 316-24.

Vilcek S, Belák S. 1996 Genetic identification of pestivirus strain Frijters as a border disease virus from pigs. J Virol Methods. 60(1): 103-8.

EXTENDED GENETIC DIVERSITY OF BOVINE VIRAL DIARRHEA VIRUS AND FREQUENCY OF GENOTYPES AND SUBTYPES IN CATTLE IN ITALY BETWEEN 1995 AND 2013

[1] T. Childs, “X Disease of Cattle – Saskatchewan”, *Canadian Journal of Comparative Medicine and Veterinary Science,* vol. 10, no. 11, pp. 316-319, 1946.

[2] P. Olafson, C.A. Mac and F. H. Fox., “An apparently new transmissible disease of cattle”, The Cornell veterinarian, vol. 36, pp. 205-213, 1946.

[3] B. Makoschey, P. T. van Gelder, V. Keijsers and D. Goovaerts, “Bovine viral diarrhoea virus antigen in foetal calf serum batches and consequences of such contamination for vaccine production”, *Biologicals*, vol. 31, no. 3, pp. 203-208, 2003.

[4] F.V. Bauermann, J.F. Ridpath, R. Weiblen and E.F. Flores, ” HoBi-like viruses: an emerging group of pestiviruses”, *Journal of Veterinary Diagnostic Investigation,* vol. 25, no. 1, pp. 6-15, 2013.

[5] S. Vilcek, B. Durkovic, M. Kolesarova, I. Greiser-Wilke and D. Paton, “Genetic diversity of international bovine viral diarrhoea virus (BVDV) isolates: identification of a new BVDV-1 genetic group”, *Veterinary Research*, vol. 35, no. 5, pp. 609–615, 2004.

[6] A. Jackova, M. Novackova, C. Pelletier, C. Audeval, E. Gueneau, A. Haffar,E. Petit, L. Rehby and S. Vilcek, “The extended genetic diversity of BVDV-1: typing of BVDV isolates from France”, *Veterinary Research Communications*, vol. 32, no. 1, pp. 7-11, 2008.

[7] K. Yeşilbağ, C. Förster, B. Bank-Wolf, Z. Yılmaz, F. Alkan, A. Ozkul, I. Burgu, S. Cedillo Rosales, H.-J. Thiel and M. König, “Genetic heterogeneity of bovine viral diarrhoea virus (BVDV) isolates from Turkey: identification of a new subgroup in BVDV-1”, *Veterinary Microbiology* , vol. 130, no. 3-4, pp. 258-67, 2008.

[8] M. Nagai, M. Hayashi, M. Itou, T. Fukutomi, H. Akashi, H. Kida, and Y. Sakoda, “Identification of new genetic subtypes of bovine viral diarrhea virus genotype 1 isolated in Japan”, *Virus Genes* , vol. 36, no. 1, pp. 135-139, 2008.

[9] F. Xue, Y. M. Zhu, J. Li, L. C. Zhu, X . G. Ren, J. K. Feng, H. F. Shi and Y.R. Gao, “Genotyping of bovine viral diarrhea viruses from cattle in China between 2005 and 2008”, *Veterinary Microbiology* , vol. 143, no. 2-4, pp. 379-383, 2010.

[10] S. Gao, J. Luo, J. Du, Y. Lang, G. Cong, J. Shao, T. Lin, F. Zhao, S. Belák, L. Liu, H. Chang and H. Yin, ”Serological and molecular evidence for natural infection of Bactrian camels with multiple subgenotypes of bovine viral diarrhea virus in Western China”, *Veterinary Microbiology* , vol. 163, no. 1-2, pp. 172-176, 2013.

[11] M. Tajima, H. R. Frey, O. Yamato, Y. Maede, V. Moennig, H. Scholz and I. Greiser-Wilke, ”Prevalence of genotypes 1 and 2 of bovine viral diarrhea virus in Lower Saxony, Germany”, *Virus Research* , vol. 76, no. 1, pp. 31-42, 2001.

[12] M. Beer, G. Wolf and O. R. Kaaden, “Phylogenetic analysis of the 5'-untranslated region of German BVDV type II isolates”, *Journal of veterinary medicine. B, Infectious diseases and veterinary public health*, vol. 49, no. 1, pp. 43-47, 2002.

[13] L. Liu, H.Xia, C. Baule and S.Belák, “Maximum likelihood and Bayesian analyses of a combined nucleotide sequence dataset for genetic characterization of a novel pestivirus, SVA/cont-08”, *Archives of Virology* , vol. 154, no. 7, pp. 1111-1116, 2009.

[14] H. Xia, B. Vijayaraghavan, S. Belák and L. Liu, ”Detection and identification of the atypical bovine pestiviruses in commercial foetal bovine serum batches”, *PLoS One* 6, e28553, 2011.

[15] K. Ståhl, J. Kampa, S. Alenius, A. Persson Wadman, C. Baule, S. Aiumlamai and S. Belák, “Natural infection of cattle with an atypical 'HoBi'-like pestivirus - implications for BVD control and for the safety of biological products”, *Veterinary Research*, vol. 38, no. 3, pp. 517-23, 2007.

[16] A. Cortez, M. B. Heinemann, A. M. M. G. Castro, R. M. Soares, A. M. V. Pinto, A. A. Alfieri, E. F. Flores, R. C. Leite and L. J. Richtzenhain, “Genetic characterization of Brazilian bovine viral diarrhea virus isolates by partial nucleotide sequencing of the 5’-UTR region”, *Pesquisa Veterinária Brasileira* , vol. 26, no. 4, pp. 211–216, 2006.

[17] E. Bianchi, M. Martins, R. Weiblen and E. F. Flores, “Perfil genotípico e antigênico de amostras do vírus da diarréia viral bovina isoladas no Rio Grande do Sul (2000-2010)”, *Pesquisa Veterinária Brasileira*, vol. 31, no. 8, pp. 649-655, 2011.

[18] N. Decaro, M. S. Lucente, V. Mari, F. Cirone, P. Cordioli, M. Camero, R. Sciarretta, M. Losurdo, E. Lorusso and C. Buonavoglia, “Atypical pestivirus and severe respiratory disease in calves, Europe”, *Emerging Infectious Diseases*, vol. 17, no. 8, pp. 1549-1552, 2011.

[19] N. Decaro, V. Mari, M. S. Lucente, R. Sciarretta, G. Elia, J. F.Ridpath and C. Buonavoglia, “Detection of a Hobi-like virus in archival samples suggests circulation of this emerging pestivirus species in Europe prior to 2007”, *Veterinary Microbiology*, vol. 167, no. 3-4, pp. 307-313, 2013.

[20] C. Bachofen, H. Stalder, U. Braun, M. Hilbe, F. Ehrensperger and E. Peterhans, “Co-existence of genetically and antigenically diverse bovine viral diarrhea viruses in an endemic situation”, *Veterinary Microbiology* , vol. 131, no. 1-2, pp. 93-102, 2008.

[21] C. Luzzago, E. Ebranati, D. Sassera, A. Lo Presti, S. Lauzi, E. Gabanelli, M. Ciccozzi and G. Zehender, "Spatial and temporal reconstruction of bovine viral diarrhea virus genotype 1 dispersion in Italy", *Infection, Genetics and Evolution*, vol. 12, no. 2, pp. 324-331, 2012.

[22] R. Strong, J. Errington, R. Cook, N. Ross-Smith, P. Wakeley and F.Steinbach, "Increased phylogenetic diversity of bovine viral diarrhoea virus type 1 isolates in England and Wales since 2001", *Veterinary Microbiology* , vol. 162, no. 2-4, pp. 315-320, 2013.

[23] C. Luzzago, C. Bandi, V. Bronzo, G. Ruffo and A. Zecconi, "Distribution pattern of bovine viral diarrhoea virus strains in intensive cattle herds in Italy", *Veterinary Microbiology* , vol. 83, no. 3, pp. 265-274, 2001.

[24] S. Vilcek, D. J. Paton, B. Durkovic, L. Strojny, G. Ibata, A. Moussa, A. Loitsch, W.Rossmanith, S. Vega, M. T. Scicluna and V. Paifi, "Bovine viral diarrhoea virus genotype 1 can be separated into at least eleven genetic groups", *Archives of Virology* , vol. 146, no. 6, pp. 99-115, 2001.

[25] E. Falcone, P. Cordioli, M. Tarantino, M. Muscillo, G. La Rosa and M. Tollis, "Genetic heterogeneity of bovine viral diarrhoea virus in Italy", *Veterinary Research Communications* , vol. 27, no. 6, pp. 485-94, 2003.

[26] S. Ciulli, E. Galletti, M. Battilani, A. Scagliarini, A. Gentile, L. Morganti and S.Prosperi, "Genetic typing of bovine viral diarrhoea virus: evidence of an increasing number of variants in Italy", *New Microbiologica*, vol. 31, no. 2, pp. 263-271, 2008.

[27] M. Giammarioli, C. Pellegrini, C. Casciari, E. Rossi and G. M. De Mia, "Genetic diversity of bovine viral diarrhea virus 1: Italian isolates clustered in at least seven subgenotypes", *Journal of Veterinary Diagnostic Investigation*, vol. 20, no. 6, pp. 783-788, 2008.

[28] V. Cannella, E. Giudice, S.Ciulli, P. Di Marco, G. Purpari, G. Cascone and A. Guercio, "Genotyping of bovine viral diarrhea viruses (BVDV) isolated from cattle in Sicily", *Comparative Clinical Pathology*, vol. 21, pp. 1733-1738, 2012.

[29] C. Letellier, P. Kerkhofs, G. Wellemans and E. Vanopdenbosch, "Detection and genotyping of bovine diarrhea virus by reverse transcription-polymerase chain amplification of the 5' untranslated region", *Veterinary Microbiology*, vol. 64, no. 2-3, pp.155-167, 1999.

[30] S. Vilcek, A. J. Herring, J. A. Herring, P. F. Nettleton, J. P. Lowings and D. J. Paton, "Pestiviruses isolated from pigs, cattle and sheep can be allocated into at least three genogroups using polymerase chain reaction and restriction endonuclease analysis", *Archives of Virology*, vol. 136, no. 3-4, pp. 309-323, 1994.

[31] D. Posada, "jModelTest: phylogenetic model averaging", *Molecular Biology and Evolution*, vol. 25, no.7, pp. 1253-1256, 2008.

[32] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar,"MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods", *Molecular Biology and Evolution*, vol. 28, no.10, pp. 2731-2739, 2011.

[33] E. Falcone, M. Tollis and G. Conti, "Bovine viral diarrhea disease associated with a contaminated vaccine", *Vaccine* , vol. 18, no. 5-6, pp. 387-388, 1999.

[34] H.W. Barkema, C.J. Bartels, L.Van Wuijckhuise, J. W. Hesselink, M. Holzhauer, M. F. Weber, P. Franken, P. A. Kock, C. J. Bruschke and G. M. Zimmer, "Outbreak of bovine virus diarrhea on Dutch dairy farms induced by a bovine herpesvirus 1 marker vaccine contaminated with bovine virus diarrhea virus type 2", *Tijdschrift voor diergeneeskunde* , vol. 126, no. 6, pp. 158-165, 2001.

[35] N. Decaro, R. Sciarretta, M.S. Lucente, V. Mari, F. Amorisco, M.L. Colaianni, P. Cordioli,A. Parisi, R. Lelli and C. Buonavoglia, "A nested PCR approach for unambiguous typing of pestiviruses infecting cattle", *Molecular and Cellular Probes,* vol. 26, no. 1, pp. 42-46, 2012.

[36] S. Peletto, F. Zuccon, M. Pitti, E. Gobbi, L.D. Marco, M. Caramelli, L. Masoero , and P.L. Acutis, " Detection and phylogenetic analysis of an atypical pestivirus, strain IZSPLV_To", *Research in Veterinary Science*, vol. 92, no. 1, pp.147-150, 2012.

[37] I. Toplak , T. Sandvik , D. Barlic-Maganja , J. Grom and D.J. Paton," Genetic typing of bovine viral diarrhoea virus: most Slovenian isolates are of genotypes 1d and 1f", *Veterinary Microbiology*, vol. 99, no. 3-4, pp. 175-185, 2004.

[38] A. Hornberg, S.R. Fernández , C. Vogl, Š. Vilček, M. Matt, M . Fink, J. Köfer and K. Schöpf, "Genetic diversity of pestivirus isolates in cattle from Western Austria", *Veterinary Microbiology*. vol. 135, no 3-4, pp.205–213, 2009.

[39] R. E. Booth, C. J. Thomas, L. M. El-Attar, G. Gunn and J. Brownlie, “A phylogenetic analysis of Bovine Viral Diarrhoea Virus (BVDV) isolates from six different regions of the UK and links to animal movement data”, *Veterinary Research* , vol. 44, no. 43, pp. 1-14, 2013.
[40] N. Decaro, V. Mari, P. Pinto, M. S. Lucente, R. Sciarretta, F. Cirone, M. L. Colaianni, G. Elia, H.-J. Thiel and C. Buonavoglia,”Hobi-like pestivirus: both biotypes isolated from a diseased animal”, *Journal of General Virology*, vol. 93, pp. 1976–1983, 2012.
[41] N. Decaro, M. S. Lucente, V. Mari, R. Sciarretta, P.Pinto, D. Buonavoglia,V. Martella and C. Buonavoglia,“Hobi-Like Pestivirus in Aborted Bovine Fetuses”, *Journal of ClinicalMicrobiology*, vol. 50, no. 2, pp. 509-512, 2012.

- **REFERENCES OF CHAPTER 2**

BAYESIAN PHYLOGEOGRAPHY OF CRIMEAN-CONGO HEMORRHAGIC FEVER VIRUS IN EUROPE

[1] Anagnostou V, Papa A (2009) Evolution of Crimean-Congo Hemorrhagic Fever virus. Infect Genet Evol 9: 948-954
[2] Schmaljohn CS, Hooper JW (2001) Bunyaviridae: the viruses and their replication. Fields virology.
[3] Ergonul O (2006) Crimean-Congo haemorrhagic fever. Lancet Infect Dis 6: 203-214.
[4] Whitehouse CA (2004) Crimean-Congo hemorrhagic fever. Antiviral Res 64: 145-160.
[5] Kondratenko VF (1976) [Importance of ixodid ticks in the transmission and preservation of the causative agent of Crimean hemorrhagic fever in foci of the infection]. Parazitologiia 10: 297-302.
[6] Hubalek Z, Rudolf I (2012) Tick-borne viruses in Europe. Parasitol Res 111: 9-36.
[7] Hoogstraal H (1981) Changing patterns of tickborne diseases in modern society. Annu Rev Entomol 26: 75-99.
[8] Leblebicioglu H (2010) Crimean-Congo haemorrhagic fever in Eurasia. Int J Antimicrob Agents 36 Suppl 1: S43-46.
[9] Swanepoel R, Shepherd AJ, Leman PA (1987) Epidemiologic and clinical features of Crimean-Congo hemorrhagic fever in southern Africa. Am J Trop Med Hyg 36: 120-132.
[10] Van De Wal BW, Joubert JR, Van Eeden PJ (1985) A nosocomial outbreak of Crimean-Congo haemorrhagic fever at Tygerberg Hospital. Part IV. Preventive and prophylactic measures. South Afr Med J 68: 729-732.
[11] Nabeth P (2007) Crimean-Congo Hemorrhagic Fever Virus. Emerging Viruses in Human Populations S0168-7069(06): 16012-16017.
[12] Bakir M, Ugurlu M, Dokuzoguz B, Bodur H, Tasyaran MA, et al. (2005) Crimean-Congo haemorrhagic fever outbreak in Middle Anatolia: a multicentre study of clinical features and outcome measures. J Med Microbiol 54: 385-389.
[13] Estrada-Pena A, Palomar AM, Santibanez P, Sanchez N, Habela MA, et al. (2012) Crimean-Congo hemorrhagic fever virus in ticks, Southwestern Europe, 2010. Emerg Infect Dis 18: 179-180.
[14] Chumakov MP, Petrova SP, Sondak VA (1945) [The encephalitides caused by viruses; experimental transmission of the virus of Japanese encephalitis to different strains of ticks]. Med Parazitol (Mosk) 14: 18-24.
[15] Papa A, Tzala E, Maltezou HC (2011) Crimean-Congo hemorrhagic fever virus, northeastern Greece. Emerg Infect Dis 17: 141-143.
[16] Jameson LJ, Ramadani N, Medlock JM (2012) Possible drivers of Crimean-Congo hemorrhagic fever virus transmission in Kosova. Vector Borne Zoonotic Dis 12: 753-757.
[17] Aradaib IE, Erickson BR, Mustafa ME, Khristova ML, Saeed NS, et al. (2010) Nosocomial outbreak of Crimean-Congo hemorrhagic fever, Sudan. Emerg Infect Dis 16: 837-839.
[18] Chinikar S, Ghiasi SM, Hewson R, Moradi M, Haeri A (2010) Crimean-Congo hemorrhagic fever in Iran and neighboring countries. J Clin Virol 47: 110-114.

[19] Gargili A, Midilli K, Ergonul O, Ergin S, Alp HG, et al. (2011) Crimean-Congo hemorrhagic fever in European part of Turkey: genetic analysis of the virus strains from ticks and a seroepidemiological study in humans. Vector Borne Zoonotic Dis 11: 747-752.
[20] Han N, Rayner S (2011) Epidemiology and mutational analysis of global strains of Crimean-Congo haemorrhagic fever virus. Virol Sin 26: 229-244.
[21] Hewson R, Chamberlain J, Mioulet V, Lloyd G, Jamil B, et al. (2004) Crimean-Congo haemorrhagic fever virus: sequence analysis of the small RNA segments from a collection of viruses world wide. Virus Res 102: 185-189.
[22] Mild M, Simon M, Albert J, Mirazimi A (2010) Towards an understanding of the migration of Crimean-Congo hemorrhagic fever virus. J Gen Virol 91: 199-207.
[23] Carroll SA, Bird BH, Rollin PE, Nichol ST (2010) Ancient common ancestry of Crimean-Congo hemorrhagic fever virus. Mol Phylogenet Evol 55: 1103-1110.
[24] Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.
[25] Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18: 502-504.
[26] Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25: 1253-1256.
[27] Guindon S, Lethiec F, Duroux P, Gascuel O (2005) PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res 33: W557-559.
[28] Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution.
[29] Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. Molecular Biology and Evolution 22: 1185-1192.
[30] Kass RE, Raftery AE (1995) Bayes factors. Journal of American Statistical Association 90: 773-795.
[31] Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. PLoS Comput Biol 5: e1000520.
[32] Bielejec F, Rambaut A, Suchard MA, Lemey P (2011) SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27: 2910-2912.
[33] Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123: 603-613.
[34] Deyde VM, Khristova ML, Rollin PE, Ksiazek TG, Nichol ST (2006) Crimean-Congo hemorrhagic fever virus genomics and global diversity. J Virol 80: 8834-8842.
[35] Ozkaya E, Dincer E, Carhan A, Uyar Y, Ertek M, et al. (2010) Molecular epidemiology of Crimean-Congo hemorrhagic fever virus in Turkey: occurrence of local topotype. Virus Res 149: 64-70.
[36] Ciccozzi M, Peletto S, Cella E, Giovanetti M, Lai A, et al. (2013) Epidemiological history and phylogeography of West Nile virus lineage 2. Infect Genet Evol 17C: 46-50.
[37] Zehender G, Ebranati E, Bernini F, Lo Presti A, Rezza G, et al. (2011) Phylogeography and epidemiological history of West Nile virus genotype 1a in Europe and the Mediterranean basin. Infect Genet Evol 11: 646-653.
[38] Zehender G, Ebranati E, Gabanelli E, Shkjezi R, Lai A, et al. (2012) Spatial and temporal dynamics of hepatitis B virus D genotype in Europe and the Mediterranean Basin. PLoS One 7: e37198.
[39] Di Nardo A, Knowles NJ, Paton DJ (2011) Combining livestock trade patterns with phylogenetics to help understand the spread of foot and mouth disease in sub-Saharan Africa, the Middle East and Southeast Asia. Rev Sci Tech 30: 63-85.
[40] Slingenbergh JI, Gilbert M, de Balogh KI, Wint W (2004) Ecological sources of zoonotic diseases. Rev Sci Tech 23: 467-484.
[41] Roeder PL, Taylor WP, Rweyemamu MM (2006) Rinderpest in the twentieth and twenty-first centuries. Academic Press.
[42] Kovalev SY, Chernykh DN, Kokorev VS, Snitkovskaya TE, Romanenko VV (2009) Origin and distribution of tick-borne encephalitis virus strains of the Siberian subtype in the Middle Urals, the north-west of Russia and the Baltic countries. J Gen Virol 90: 2884-2892.

[43] Shepherd AJ, Swanepoel R, Leman PA, Shepherd SP (1987) Field and laboratory investigation of Crimean-Congo haemorrhagic fever virus (Nairovirus, family Bunyaviridae) infection in birds. Trans R Soc Trop Med Hyg 81: 1004-1007.
[44] Anonymous (2010) EFSA Panel on Animal and Welfare-Scientific Opinion on the Role of Tick Vectors in the Epidemiology of Crimean Congo Hemorrhagic Fever and African Swine Fever in Eurasia. EFSA Journal 8: 1703.
[45] Mahzounieh M, Dincer E, Faraji A, Akin H, Akkutay AZ, et al. (2012) Relationship between Crimean-Congo hemorrhagic fever virus strains circulating in Iran and Turkey: possibilities for transborder transmission. Vector Borne Zoonotic Dis 12: 782-785.
[46] Anonymous (2012) EFSA Panel on Animal Health and Welfare-Scientific Opinion on foot-and-mouth disease in Thrace. EFSA Journal 10: 2635.
[47] Sherman DM (2011) The spread of pathogens through trade in small ruminants and their products. Rev Sci Tech 30: 207-217.
[48] Midilli K, Gargili A, Ergonul O, Sengoz G, Ozturk R, et al. (2007) Imported Crimean-Congo hemorrhagic fever cases in Istanbul. BMC Infect Dis 7: 54.
[49] Kunchev A, Kojouharova M (2008) Probable cases of Crimean-Congo-haemorrhagic fever in Bulgaria: a preliminary report. Euro Surveill 13.
[50] Papa A, Maltezou HC, Tsiodras S, Dalla VG, Papadimitriou T, et al. (2008) A case of Crimean-Congo haemorrhagic fever in Greece, June 2008. Euro Surveill 13.

• REFERENCES OF CHAPTER 3

PHYLOGEOGRAPHY AND PHYLODYNAMICS OF EUROPEAN GENOTYPE 3 HEPATITIS E VIRUS

Bouamra, Y., et al., 2014. Emergence of autochthonous infections with hepatitis E virus of genotype 4 in Europe. Intervirology 57, 43-48.

Ciccozzi, M., et al., 2011. Reconstruction of the evolutionary dynamics of the hepatitis C virus 1b epidemic in Turkey. Infect Genet Evol 11, 863-868.

Dalton, H.R., et al., 2008. Hepatitis E: an emerging infection in developed countries. The Lancet infectious diseases 8, 698-709.

Drummond, A.J., et al., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29, 1969-1973.

Edwards, C.J., et al., 2011. Ancient hybridization and an Irish origin for the modern polar bear matriline. Current biology : CB 21, 1251-1258.

Garbuglia, A.R., et al., 2013. Hepatitis E virus genotype 4 outbreak, Italy, 2011. Emerging infectious diseases 19, 110-114.

Hoofnagle, J.H., et al., 2012. Hepatitis E. The New England journal of medicine 367, 1237-1244.

Inoue, J., et al., 2006. Analysis of human and swine hepatitis E virus (HEV) isolates of genotype 3 in Japan that are only 81-83 % similar to reported HEV isolates of the same genotype over the entire genome. The Journal of general virology 87, 2363-2369.

Johne, R., et al., 2010. Detection of a novel hepatitis E-like virus in faeces of wild rats using a nested broad-spectrum RT-PCR. The Journal of general virology 91, 750-758.

Kamar, N., et al., 2012. Hepatitis E. Lancet 379, 2477-2488.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of American Statistical Association 90, 773-795.

Kuniholm, M.H., et al., 2009. Epidemiology of hepatitis E virus in the United States: results from the Third National Health and Nutrition Examination Survey, 1988-1994. The Journal of infectious diseases 200, 48-56.

Lemey, P., et al., 2009. Bayesian phylogeography finds its roots. PLoS Comput Biol 5, e1000520.

Lu, L., et al., 2006. Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. Reviews in medical virology 16, 5-36.

Mansuy, J.M., et al., 2008. High prevalence of anti-hepatitis E virus antibodies in blood donors from South West France. Journal of medical virology 80, 289-293.

Meng, X.J., 2011. From barnyard to food table: the omnipresence of hepatitis E virus and risk for zoonotic infection and food safety. Virus research 161, 23-30.
Nakano, T., et al., 2012a. Molecular epidemiology and genetic history of European-type genotype 3 hepatitis E virus indigenized in the central region of Japan. Infect Genet Evol 12, 1524-1534.
Nakano, T., et al., 2012b. New findings regarding the epidemic history and population dynamics of Japan-indigenous genotype 3 hepatitis E virus inferred by molecular evolution. Liver international : official journal of the International Association for the Study of the Liver 32, 675-688.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol Biol Evol 25, 1253-1256.
Pujols, J., et al., 2011. Commercial spray-dried porcine plasma does not transmit porcine circovirus type 2 in weaned pigs challenged with porcine reproductive and respiratory syndrome virus. Veterinary journal (London, England : 1997) 190, e16-20.
Purdy, M.A., Khudyakov, Y.E., 2010. Evolutionary history and population dynamics of hepatitis E virus. PloS one 5, e14376.
Pybus, O.G., et al., 2000. An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics 155, 1429-1437.
Rodriguez, F., et al., 1990. The general stochastic model of nucleotide substitution. J Theor Biol 142, 485-501.
Romano, L., et al., 2013. Hepatitis E: a puzzling double-faced disease. Annali di igiene : medicina preventiva e di comunita 25, 169-180.
Schmidt, H.A., et al., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18, 502-504.
Slatkin, M., Maddison, W.P., 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123, 603-613.
Strimmer, K., von Haeseler, A., 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc Natl Acad Sci U S A 94, 6815-6819.
Suchard, M.A., et al., 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol Biol Evol 18, 1001-1013.
Takahashi, M., et al., 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, 902-908.
Tanaka, Y., et al., 2006. Molecular tracing of Japan-indigenous hepatitis E viruses. The Journal of general virology 87, 949-954.
Zehender, G., et al., 2012. Spatial and temporal dynamics of hepatitis B virus D genotype in Europe and the Mediterranean Basin. PloS one 7, e37198.
Zehender, G., et al., 2008. Different evolutionary rates and epidemic growth of hepatitis B virus genotypes A and D. Virology 380, 84-90.
Zhao, C., et al., 2009. A novel genotype of hepatitis E virus prevalent among farmed rabbits in China. Journal of medical virology 81, 1371-1379.

ANNEX
SUPPORTING INFORMATION

SUPPORTING MATERIALS ARE LISTED BY CHAPTER:

CHAPTER 1

Spatial and temporal phylogeny of border disease virus in pyrenean chamois

- Figure S1 Likelihood map of the 95 BDV 5'UTR

Extended genetic diversity of bovine viral diarrhea virus and frequency of genotypes and subtypes in cattle in italy between 1995 and 2013

- Figure S1. Phylogenetic tree based on the 5'-UTR of 371 Italian sequences representative of all BVDV genotypes and subtypes detected between 1995 and 2013 and reference BVDV-1, BVDV-2, and HoBi-like strains
- Table S1. Temporal distribution of BVDV genotypes and subtypes in Italy.

CHAPTER 2

Bayesian phylogeography of Crimean-Congo Hemorrhagic Fever virus in Europe

- Figure S1 Evaluation of the impact of sampling heterogeneity on the phylogeographic reconstruction.
- Figure S2 Likelihood map of the 121 CCHFV S gene sequences
- Figure S3 Maximum likelihood tree of the 121 CCHFV S gene sequences. The numbers on the branches represent bootstrap values (see Materials and Methods for details). The previously described viral genotypes have been highlighted.
- Table S1 Accession numbers and characteristics of the CCHFV sequences used in the study.

CHAPTER 3

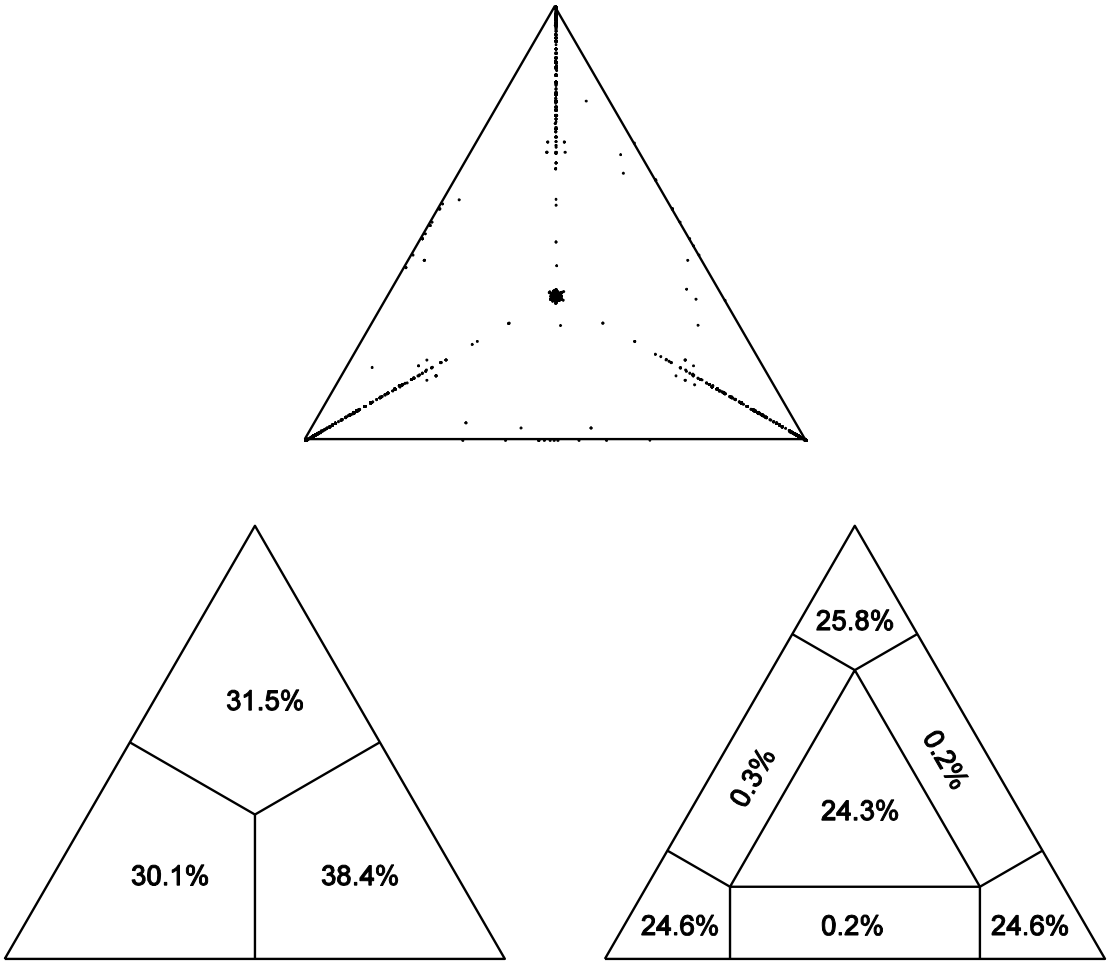
Phylogeography and phylodynamics of european genotype 3 hepatitis e virus

- Figure 1S Likelihood mapping of 208 HEV-3 ORF-2 sequences
- Figure 2S Maximum parsimony migration patterns of HEV sequences to/from different geographical groups.
- Table 1S. List of the isolates used in the analysis, with sequence accession numbers, sampling locations and years (when available), and host species.
- Table 2S. Comparison between the tMRCAs and credibility intervals (95%HPD) of the main clades estimated by using the external (Nakano et al., 2012b) or the internal (heterochronous sequences) calibration approaches.

CHAPTER 1

Spatial and temporal phylogeny of border disease virus in pyrenean chamois

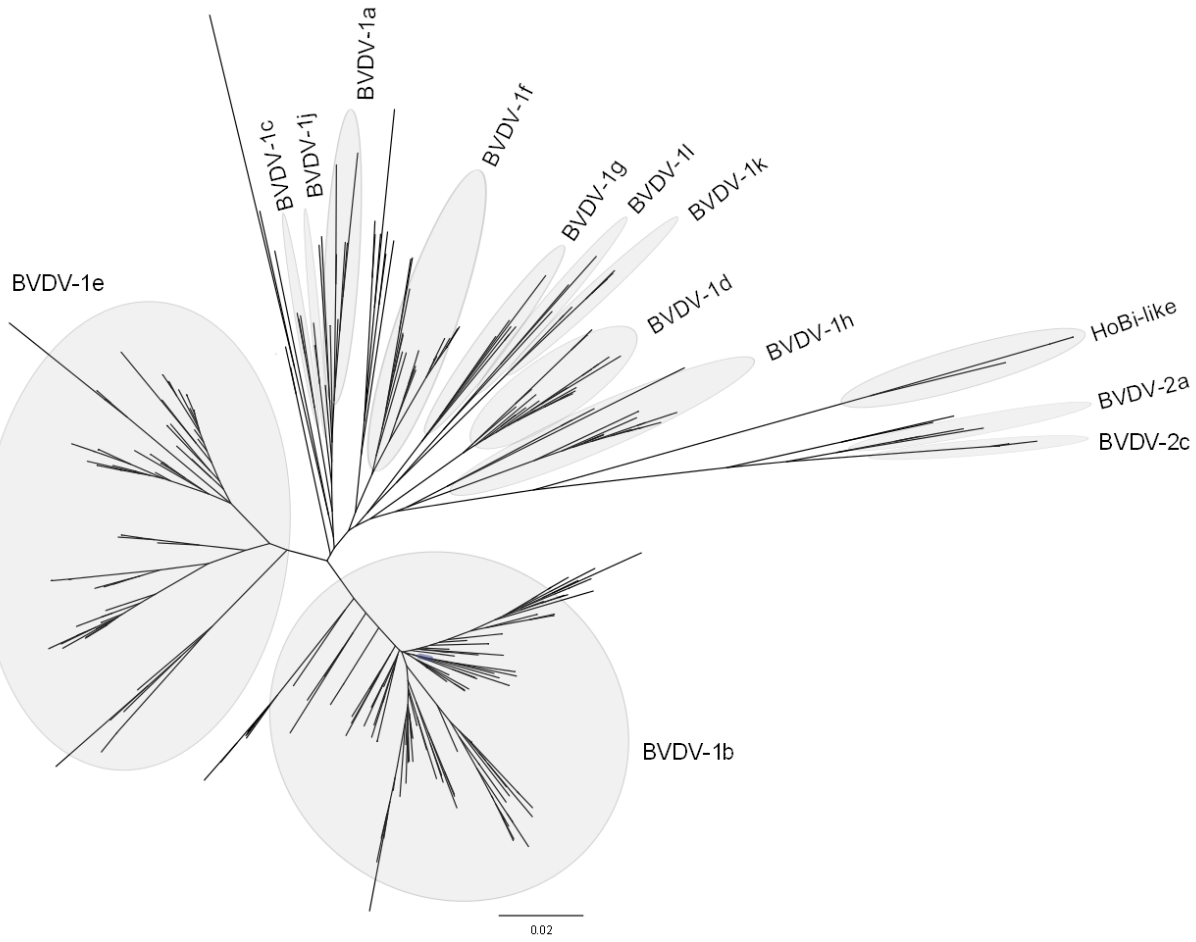
Figure S1 Likelihood map of the 95 BDV 5'UTR sequences. Each dot represents the likelihoods of the three possible unrooted trees per quartet randomly selected from the data set: the dots near the corners or sides respectively represent tree-like (fully resolved phylogenies in which one tree is clearly better than the others) or network-like phylogenetic signals (three regions in which it is not possible to decide between two topologies). The central area of the map represents a star-like signal (the region in which the star tree isoptimal tree). The numbers indicate the percentage of dots distribution.



CHAPTER 1

Extended genetic diversity of bovine viral diarrhea virus and frequency of genotypes and subtypes in cattle in italy between 1995 and 2013

Figure S1. Phylogenetic tree based on the 5'-UTR of 371 Italian sequences representative of all BVDV genotypes and subtypes detected between 1995 and 2013 and reference BVDV-1, BVDV-2, and HoBi-like strains. Molecular evolutionary genetics analyses were performed with MEGA5 using the neighbor-joining method. The genotypes and subtypes detected in Italy have been highlighted.



CHAPTER 1

Extended genetic diversity of bovine viral diarrhea virus and frequency of genotypes and subtypes in cattle in italy between 1995 and 2013

Table S1. Temporal distribution of BVDV genotypes and subtypes in Italy.

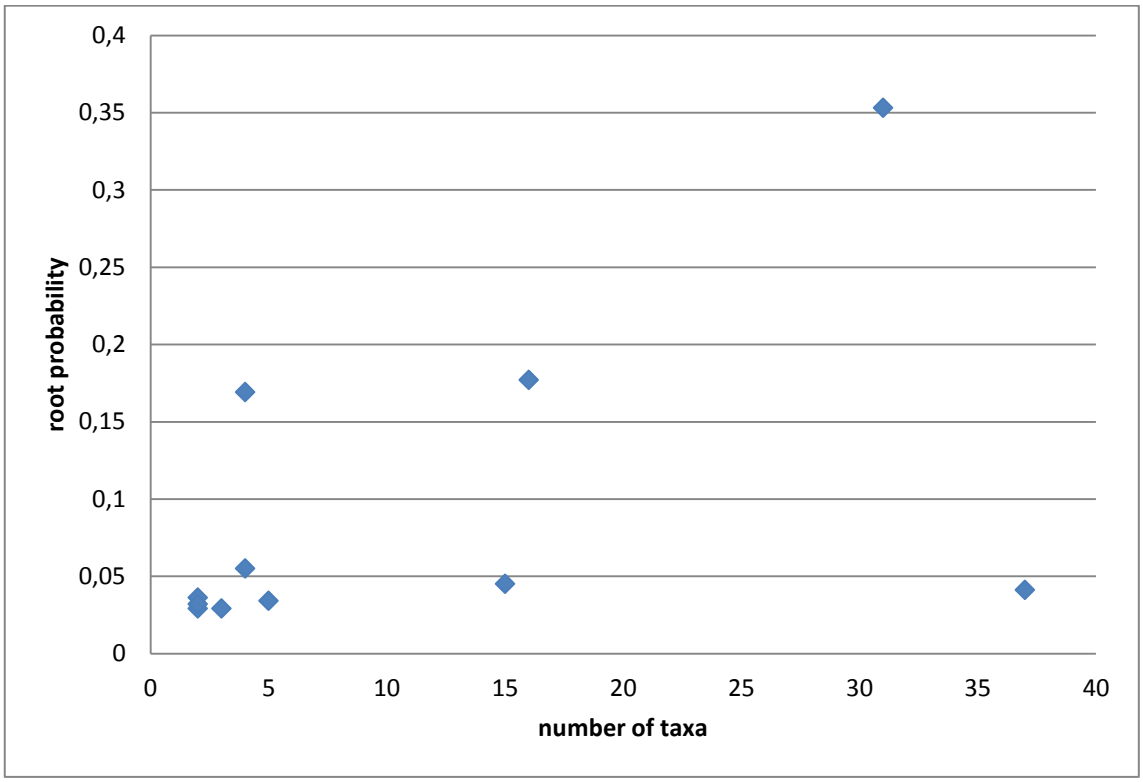
Year	BVDV genotype and subtype*												
	1a	1b	1c	1d	1e	1f	1g	1h	1j	1k	1l	2	HoBi-like
1995		■		■					■			■	
1996		■			■			■					
1997		■			■							■	
1998		■											
1999		■		■	■	■		■				■	
2000	■	■			■								
2001		■			■					■		■	
2002		■		■	■	■	■	■		■		■	
2003		■			■		■	■					
2004		■			■			■				■	
2005	■	■		■	■			■					
2006		■		■	■	■		■					
2007	■	■		■	■	■		■			■		■
2008		■			■	■							
2009	■	■		■	■	■		■					
2010		■	■	■	■	■	■	■					■
2011		■			■	■		■		■			■
2012		■			■	■							
2013		■			■								

* colored boxes indicate detection of the BVDV genotype and subtype.

CHAPTER 2

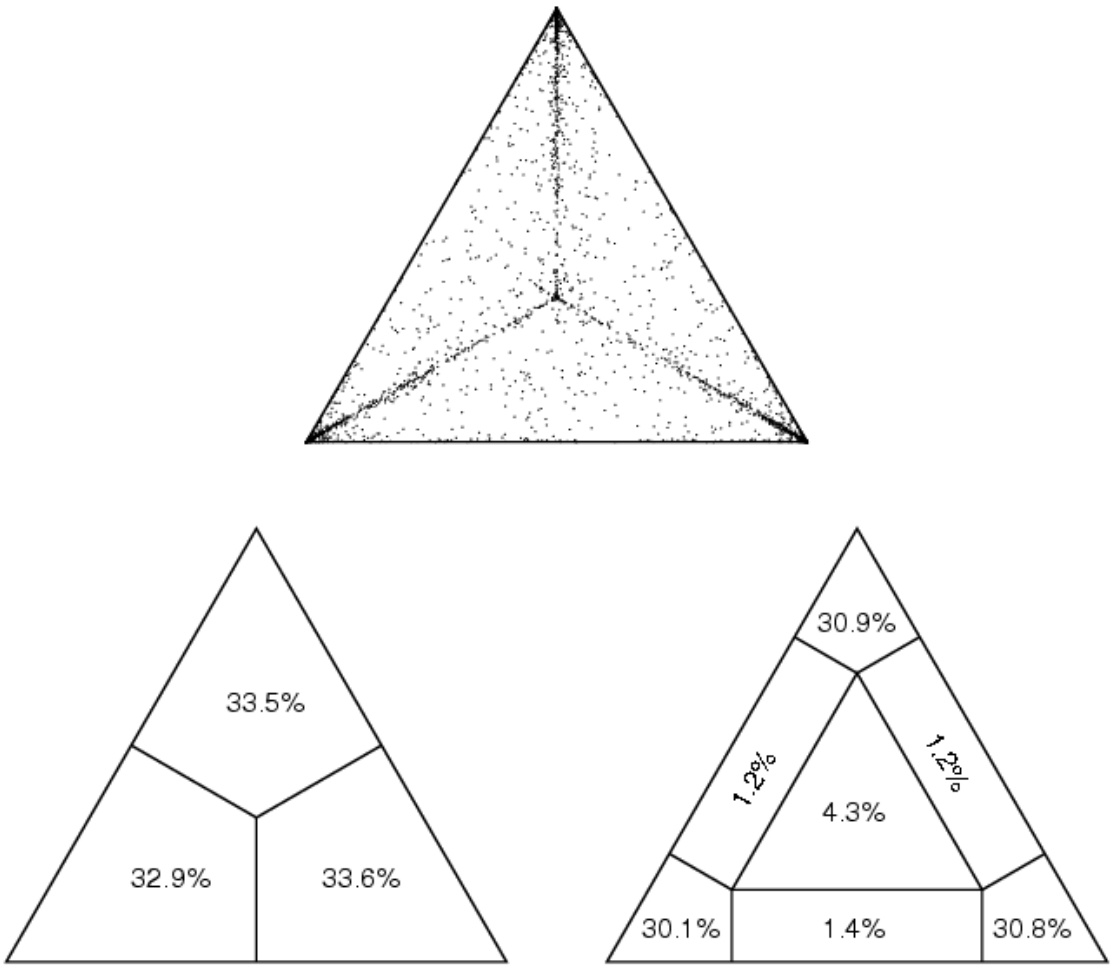
Bayesian phylogeography of Crimean-Congo Hemorrhagic Fever virus in Europe

Figure S1 Evaluation of the impact of sampling heterogeneity on the phylogeographic reconstruction. The figure shows the root state probability as a function of the location sample size. Randomisation analysis of the tip-localities throughout the MCMC analysis revealed a low level of correlation between the number of taxa per locality and the root-location probability



Bayesian phylogeography of Crimean-Congo Hemorrhagic Fever virus in Europe

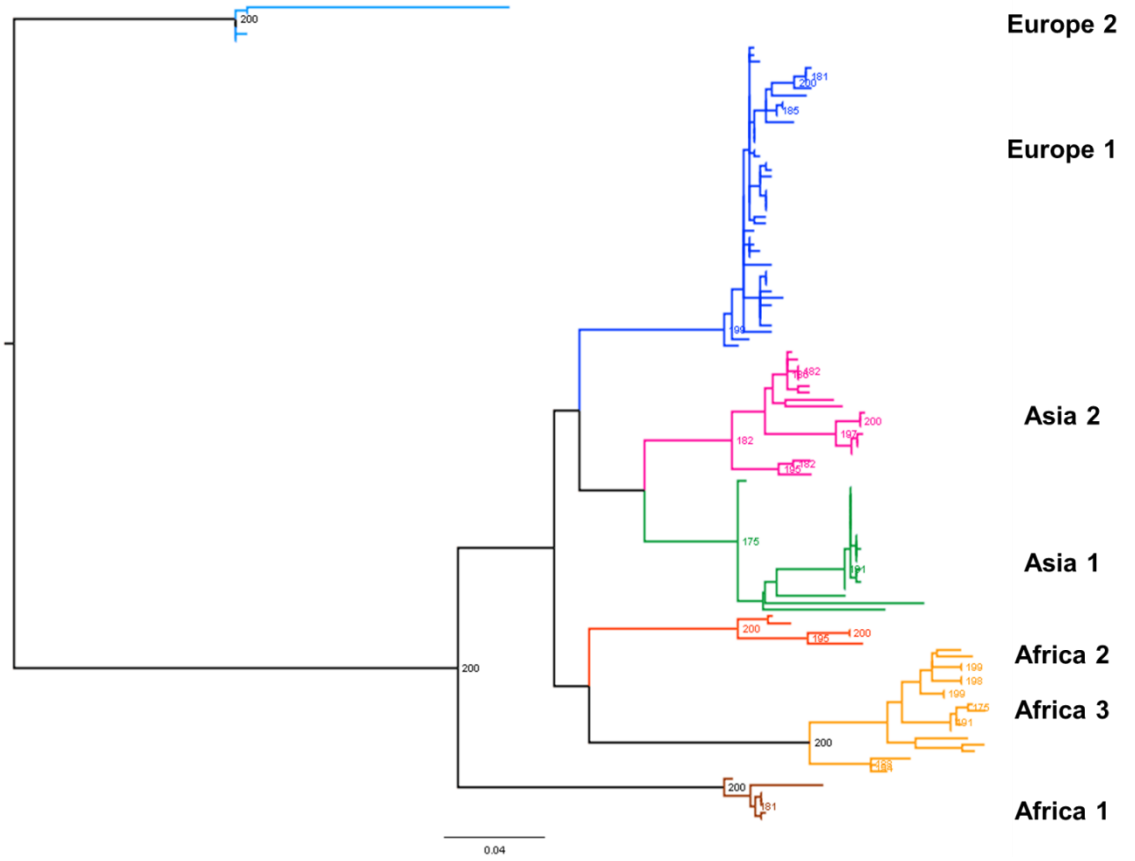
Figure S2 Likelihood map of the 121 CCHFV S gene sequences. Each dot represents the likelihoods of the three possible unrooted trees per quartet randomly selected from the data set: the dots near the corners and sides respectively represent tree-like (fully resolved phylogenies in which one tree is clearly better than the others) and network-like phylogenetic signals (three regions in which it is not possible to decide between two topologies). The central area of the map represents a star-like signal (the region in which the star tree is the optimal tree). The numbers indicate the percentage of dots in the centre of the triangle.



CHAPTER 2

Bayesian phylogeography of Crimean-Congo Hemorrhagic Fever virus in Europe

Figure S3 Maximum likelihood tree of the 121 CCHFV S gene sequences. The numbers on the branches represent bootstrap values (see Materials and Methods for details). The previously described viral genotypes [22] have been highlighted.



CHAPTER 2

Bayesian phylogeography of Crimean-Congo Hemorrhagic Fever virus in Europe

Table S1- Accession numbers and characteristics of the CCHFV sequences used in the study.

Sample	Accession number	Isolate	Sampling Location	Host	Sampling Year
AP92@75	U04958	AP92	Greece	Rhipicephalus bursa	1975
SE1@93	U15091	ArD 97268	Senegal	H. truncatum	1993
SE2@93	U15090	ArD 97264	Senegal	H. marginatum rufipes	1993
SE3@72	DQ211640	ArD15786	Senegal	goat	1972
SE4@69	DQ211639	ArD8194	Senegal	H. truncatum	1969
SE5@69	U88411	DAK 8194	Senegal		1969
SE6@69	U15021	ArD 8194	Senegal		1969
UGA1@56	U88416	UGANDA 3010	Uganda	human	1956
UGA2@81	DQ076415	SPU128/81/7	Uganda	Hyalomma (tick)	1981
UGA3@58	DQ076413	Semunya	Uganda		1958
CAR1@73	U15092	ArB 604	CAR	H.nitidum(tick)	1973
NIG1@66	U88410	IbAr10200	Nigeria	H. excavatum (tick)	1966
MAD1@85	U15024	ArMg 951	Madagascar	Boophilus microplus (tick)	1985
SA1@86	U84638	SPU 422/86	South Africa		1986
SA2@87	DQ211647	SPU103/87	South Africa	human	1987
SA3@85	DQ211646	SPU97/85	South Africa	human	1985
SA4@89	FJ435421	SPU 337/89	South Africa	human	1989
SA5@88	FJ435422	SPU 45/88	South Africa	human	1988
SA6@88	FJ435423	SPU 71/88	South Africa	human	1988

SA7@81	DQ076416	SPU4/81	South Africa	human	1981
SA8@85	U84636	SPU 247/85	South Africa		1985
SA9@86	U84635	SPU 582/86	South Africa		1986
SA10@88	U84637	45/88	South Africa		1988
SA11@85	DQ211648	SPU415/85	South Africa	human	1985
MAU1@84	U15089	ArD 39554	Mauritania	H. marginatum rufipes	1984
MAU2@84	DQ211641	ArD39554	Mauritania	H. marginatum rufipes	1984
MAU3@88	U15023	HD 49199	Mauritania	human	1988
CO1@56	DQ211650	UG3010	Congo	human	1956
SUD1@08	GQ862371	Sudan Al-Fulah 3-2008	Sudan	human	2008
SUD2@08	GQ862372	Sudan Al-Fulah 4-2008	Sudan	human	2008
SUD3@09	HQ378179	Sudan AB1	Sudan	human	2009
BF1@83	U15093	HD 38562	Burkina Faso	human	1983
OMN1@97	DQ211645	Oman strain	Oman		1997
AL5@04	KC846094	178/04	Albania	human	2004
AL6@03	KC846093	Albania 23/03	Albania	human	2003
KO1@00	AF404507	Kosovo strain	Kosovo	human	2000
KO2@01	AF428144	9553/01	Kosovo	human	2001
KO4@01	DQ133507	Kosovo Hoti	Kosovo	human	2001
BU9@08	FJ472634	Bul2008	Bulgaria	human	2008
BU10@78	AY277676	Bul/hU517	Bulgaria	human	1978
T1@06	DQ983741	CTF-Hu7/06	Turkey	human	2006
T2@06	GQ337053	Turkey-Kelkit06	Turkey	human	2005
T3@08	FJ392604	KMAG-Hu-08-01	Turkey	human	2008
T4@07	EU727456	GOU-OT07	Turkey	human	2007
T5@03	DQ211649	Turkey200310849	Turkey	human	2003

T6@05	EF432649	Kelkit S/Turkey-hu11/2005	Turkey	human	2005
T7@04	EF432640	Kelkit S/Turkey-hu1/2004	Turkey	human	2004
T8@08	FJ601895	1503	Turkey	human	2008
T9@08	FJ601863	BOLU173-2006-201	Turkey	human	2008
Tke10@04	EF432639	hu2/2004	Turkey	human	2004
Tke11@04	EF432653	hu15/2004	Turkey	human	2004
Tke12@05	EF432650	hu12/2005	Turkey	human	2005
Tke13@05	EF432647	hu9/2005	Turkey	human	2005
Tky19@09	HQ188915	Kayseri 3159-2009	Turkey	human	2009
Tay20@10	HQ188918	Aydin 193-2010	Turkey	human	2010
Tay21@09	HQ173901	Aydin 919-2009	Turkey	human	2009
Tyo22@10	HQ188920	Yozgat 714-2010	Turkey	human	2010
Tto24@06	GU324990	k16-2-kk536	Turkey	tick	2006
Tto25@06	GU324991	k36-kk536	Turkey	tick	2006
Tto26@06	GU324992	k52-kk536R	Turkey	tick	2006
Tto27@06	GU324993	k53-kk536	Turkey	tick	2006
Tto28@09	HQ173902	Tokat 1728-2009	Turkey	human	2009
Tar29@10	HQ675002	Artvin 1913-2010	Turkey	human	2010
Tba30@09	HQ821873	BAYBURT 2946-2009	Turkey	human	2009
Tca31@09	HQ675006	Cankiri 1079-2009	Turkey	human	2009
Tco32@09	HQ675010	Corum 4319-2009	Turkey	human	2009
Ter33@10	HQ664913	Erzurum 1912-2010	Turkey	human	2010
Tez34@10	HQ675008	Erzincan 1910-2010	Turkey	human	2010
Tka35@10	HQ675007	Karabuk 590-2010	Turkey	human	2010
Tks36@09	HQ173895 -	Kastamonu 1420-2009	Turkey	human	2009

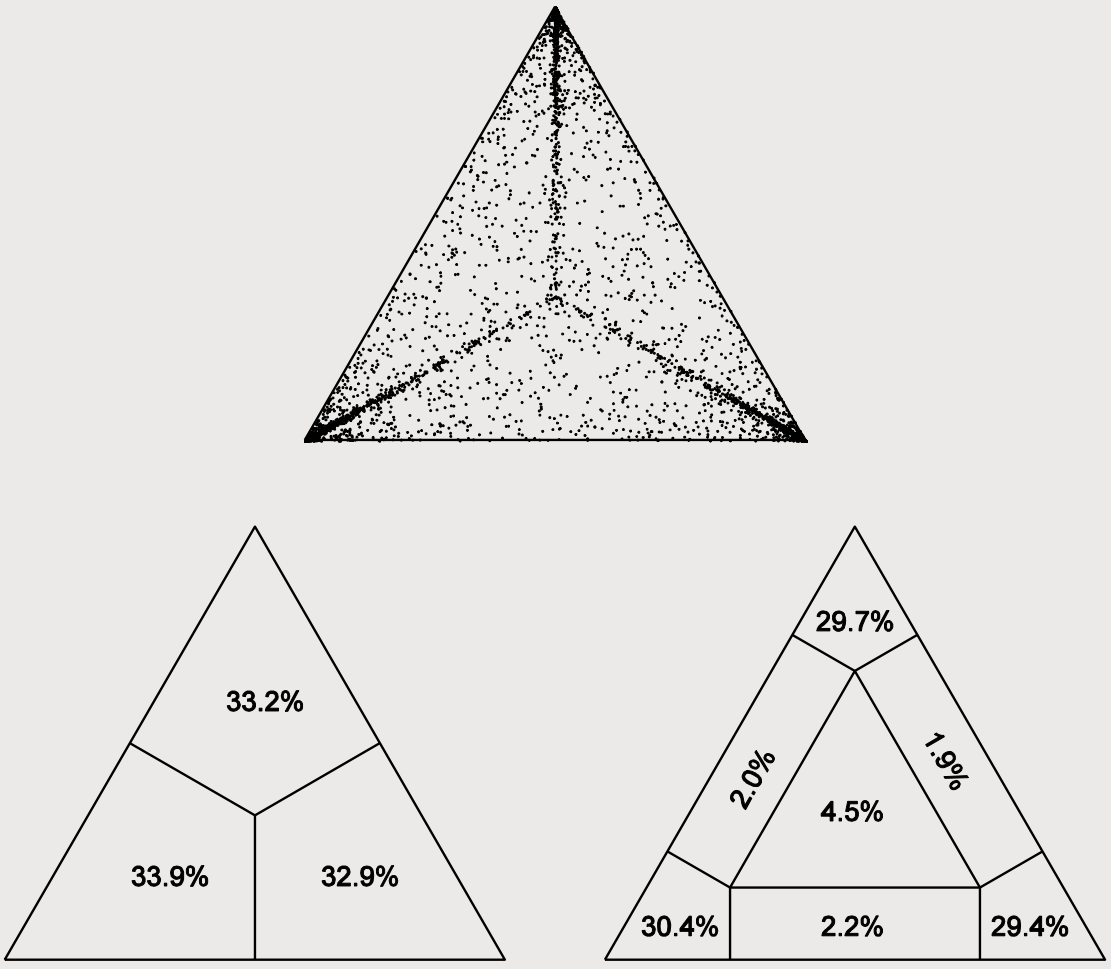
Tsa37@09	HQ675003	Samsun 1030-2009	Turkey	human	2009
Tsi38@09	HQ173893	Sivas 405-2009	Turkey	human	2009
Tel39@06	DQ983785	ZAM57/06	Turkey	H.marginatum	2006
I1@78	U15022	ArTeh 193-3	Iran		1978
Isi2@02	AY366373	766/02	Iran	human	2002
Isi3@02	AY366374	756/02	Iran	human	2002
Isi4@02	AY366376	714/02	Iran	human	2002
Isi5@02	AY366377	782/02	Iran	human	2002
Isi6@02	AY366378	786/02	Iran	human	2002
Isi7@02	AY366379 -	787/02	Iran	human	2002
Iha8@07	GU456723	CT9	Iran	Rhipicephalus sanguineus	2007
Iha9@07	GU456724	CT10	Iran	H. detritum (tick)	2007
Iha10@07	GU456725	CT12	Iran	H. detritum (tick)	2007
Iha11@07	GU456726	CT13	Iran	Argas reflexus (tick)	2007
Iha12@07	GU456727	CT14	Iran	H. anatolicum (tick)	2007
Iha13@07	GU456728	CT15	Iran	H. detritum (tick)	2007
Iqo14@02	AY366375	729/02	Iran	human	2002
IRQ1@79	AJ538196	Baghdad 12	Iraq	human	1979
PAK1@65	U88414	JD 206	Pakistan	H. anatolicum	1965
PAK2@76	AF527810	Matin	Pakistan		1976
PAK4@00	AJ538198	SR3	Pakistan	human	2000
AFG1@09	HM452305	Afg09	Afghanistan	human	2009
UZB1@85	AF481799	Uzbek/TI10145	Uzbekistan	H. asiaticum	1985
UZB2@67	AY223475	Hodzha	Uzbekistan	human	1967
TAJ1@90	AY049083	TAJ/HU8966	Tajikistan	human	1990
TAJ3@91	AY297691	TAJ/HU8978	Tajikistan	human	1991
R1@67	DQ211644	Kashmanov	Russia	human	1967
Rst2@00	AF481802	STV/HU29223 Russia	Stavropol	human	2000

Ras6@67	DQ211643	Drosdov	Astrakhan	human	1967
Rro15@00	AY277672	ROS/TI28044 ROS/HUVLV-100	Rostov	H. marginatum	2000
Rro16@02	DQ206447	Russia	Rostov	human	2002
CH8@04	GU477494	79121M18	China	tick	2004
CH9@04	FJ562093	YL04057	China	tick	2004
CH10@05	DQ227496	CLT/TI05146	China	tick	2005
CH11@05	DQ227495	CYT/TI05099	China	tick	2005
CH12@05	DQ217602	CYL/TI05035	China		2005
CH13@68	M86625	C68031	China		1968
CH14@70	AF415236	c7001	China		1970
CH15@79	AF358784	c79121	China		1979
CH16@66	AJ010648	c66019	China	human	1966
CH17@88	AY029157	c88166	China		1988
CH18@84	AJ010649	c8402	China	H. asiaticum	1984
CH19@68	DQ211642	C-68031	China	sheep	1968
CH20@75	AF362080	c75024	China		1975
CH21@78	AF354296	c7803	China		1978
CH22@68	U88413	HY 13	China	H. asiaticum	1968
T40@08	FJ392604	KMAG-Hu-08-01	Turkey	human	2008
T41@08	FJ392601	Tr-T-03	Turkey	Rhipicephalus bursa	2008
T42@08	FJ392603	Tr-T-11	Turkey	Rhipicephalus bursa	2008
Tis23@07	EU057975	KMAG-Hu-07-01	Turkey	human	2008
G1@08	EU871766	66/08-Rodopi	Greece	human	2007 2008

CHAPTER 3

Phylogeography and phylodynamics of european genotype 3 hepatitis e virus

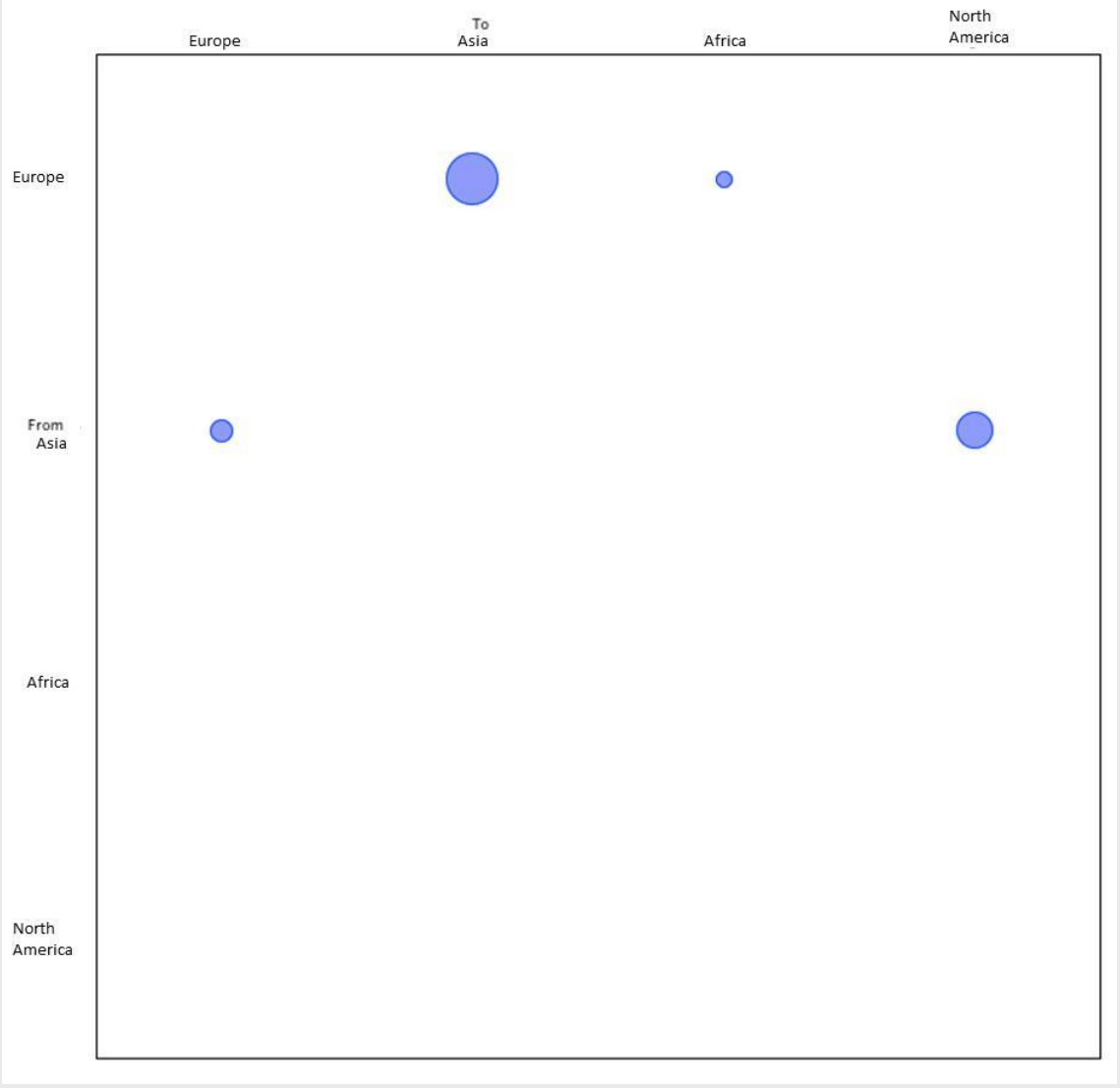
Figure 1S Likelihood mapping of 208 HEV-3 ORF-2 sequences. Each dot represents the likelihoods of the three possible unrooted trees per quartet randomly selected from the data set : the dots near the corners or the sides respectively represent tree-like (fully resolved phylogenies where one tree is clearly better than the others) or network-like phylogenetic signals (three regions in which it is not possible to decide between two topologies). The central area of the likelihood map represents a star-like signal (the region where the star tree is the optimal tree). The numbers indicate the percentage of dots.



CHAPTER 3

Phylogeography and phylodynamics of european genotype 3 hepatitis e virus

Figure 2S Maximum parsimony migration patterns of HEV sequences to/from different geographical groups. The bubblegram shows the frequency of gene flow (migrations) to/from different locations as percentages of the total number of observed migrations estimated from the maximum likelihood trees for the different subtypes using a modified version of the Slatkin and Maddison test. The surface of each circle is proportional to the percentage of observed migrations written within the circle.



CHAPTER 3

Phylogeography and phylodynamics of european genotype 3 hepatitis e virus

Table 1S. List of the isolates used in the analysis, with sequence accession numbers, sampling locations and years (when available), and host species.

ACCESSION NUMBER	CODE	LOCATION	ORIGIN
HM446627(HU IT-2)	44IHS@09	Europe	homo sapiens
HM446629(HU IT-4)	46IHS@03	Europe	homo sapiens
HM446628(HU IT-3)	45IHS@07	Europe	homo sapiens
HM446630(HU IT-5)	47IHS@05	Europe	homo sapiens
HM446631(HU IT-6)	48IHS@09	Europe	homo sapiens
EF053274	59F@06	Europe	
EU543566	57FHM@07	Europe	homo sapiens
FJ464744	54SHM@08	Europe	homo sapiens
HM623777	49ASW@08	Europe	pig
HM623776	50ASW@07	Europe	pig
HM623775	51ASW@07	Europe	pig
HM066937	52FHM@09	Europe	homo sapiens
HM439250	56CHM@09	Asia	homo sapiens
FJ600536	55OSW@08	Africa	pig
GQ398024	53FPI@09	Europe	pig
AB073912	58JHS@03	Asia	homo sapiens
AB443627	1JSW@00	Asia	pig
AB074918	6JSW@02	Asia	pig
AB074918	2JHS@01	Asia	homo sapiens
AB074920	3JHS@01	Asia	homo sapiens
AB089824	4JHS@93	Asia	homo sapiens
AB091394	5JHS@98	Asia	homo sapiens
AB189071	7JDE@03	Asia	deer
AB189072	8JHS@03	Asia	homo sapiens
AB189074	9JHS@03	Asia	homo sapiens
AB189075	10JHS@03	Asia	homo sapiens

AB222182	11JHS@03	Asia	homo sapiens
AB222183	12JBO@03	Asia	boar
AB222184	13JBO@04	Asia	boar
AB290312	14JBO@05	Asia	boar
AB290313	15JHS@01	Asia	homo sapiens
AB291951	16MPI@06	Asia	pig
AB291957	17JHS@97	Asia	homo sapiens
AB291961	18JHS@04	Asia	homo sapiens
AB291962	19JHS@04	Asia	homo sapiens
AB291963	20JHS@03	Asia	homo sapiens
AB301710	21JHS@05	Asia	homo sapiens
AB369689	22JCC@03	Asia	FECAL specimens in a cell culture system
AB437316	23JHS@04	Asia	homo sapiens
AB437317	24JHS@03	Asia	homo sapiens
AB437318	25JHS@03	Asia	homo sapiens
FJ998015	26JHS@03	Asia	homo sapiens
EU375463	27GSW@07	Europe	pig
EU723512	28TPI@07	Asia	pig
EU723514	29SSW@05	Europe	pig
EU723516	30SSW@06	Europe	pig
FJ426403	31SSW@06	Europe	pig
FJ426404	32KSW@07	Asia	pig
FJ527832	33KSW@07	Asia	pig
FJ705359	34CSW@08	Asia	pig
FJ956757	35JHS@01	Asia	homo sapiens
FJ998008	38GHS@05	Europe	homo sapiens
FJ653660	39GSW@07	Europe	pig
AB248520	40THS@08	Asia	homo sapiens
AB248521	41JHS@04	Asia	homo sapiens
AB248522	42JPI@04	Asia	pig
325176962	43JPI@04	Asia	pig

372001242	63IHS@08	Europe	homo sapiens
372001238	60ISW@10	Europe	pig
372001240	61ISW@10	Europe	pig
312458542	64ISW@10	Europe	pig
312458538	65ISW@08	Europe	pig
312458534	66ISW@08	Europe	pig
312458530	67ISW@08	Europe	pig
312458526	68ISW@08	Europe	pig
312458522	69ISW@08	Europe	pig
312458518	70ISW@08	Europe	pig
312458540	71ISW@08	Europe	pig
312458536	72ISW@08	Europe	pig
312458532	73ISW@08	Europe	pig
312458528	74ISW@08	Europe	pig
312458524	75ISW@08	Europe	pig
312458520	76ISW@08	Europe	pig
312458516	77ISW@08	Europe	pig
289429490	78ISW@08	Europe	pig
289429490	79ISW	Europe	pig
AB073912	80JSW	Asia	pig
AB189070	81JBO@04	Asia	boar
AB236320	82JM@02	Asia	mongoose
AB246676	83J	Asia	
AB291952	85JHU@05	Asia	homo sapiens
AB291953	86JHU@05	Asia	homo sapiens
AB291954	87JHU@04	Asia	homo sapiens
AB291955	88JHU@06	Asia	homo sapiens
AB291956	89JHU@04	Asia	homo sapiens
AB291958	90JHU@04	Asia	homo sapiens
AB291960	91JHU@06	Asia	homo sapiens
AB369687	97JHU@98	Asia	homo sapiens

AB369691	98JHU@05	Asia	homo sapiens
AB437319	101JH@03	Asia	homo sapiens
AB443623	102JS@02	Asia	pig
AB443624	103JS@02	Asia	pig
AB443625	104JS@02	Asia	pig
AB443626	105JS@02	Asia	pig
AF060668	106UHU	United States	homo sapiens
AF060669	107UHU	United States	homo sapiens
AF082843	108USW	United States	pig
AF455784	109YS@88	Asia	pig
AP003430	110JHU	Asia	homo sapiens
AY115488	111CAHU	United States	homo sapiens
EU360977	112SESW	Europe	pig
EU495148	113FHU	Europe	homo sapiens
EU723513	114SSW	Europe	pig
EU723515	115SSW	Europe	pig
FJ998019	116GSW	Europe	pig
AB073910	117JPIb	Asia	pig
AB073911	118JPIe	Asia	pig
AB079763	119JHSb	Asia	homo sapiens
AB080579	120JHSa	Asia	homo sapiens
AB082560	121JHSa	Asia	homo sapiens
AB08256	122JHSb	Asia	homo sapiens
AB082562	123JHSb	Asia	homo sapiens
AB082563	124JHSa	Asia	homo sapiens
AB082564	125JHSa	Asia	homo sapiens
AB082565	126JHSb	Asia	homo sapiens
AB082567	127JHSb	Asia	homo sapiens
AB088418	128JHSb	Asia	homo sapiens
AB093535	129JHSe	Asia	homo sapiens
AB094203	130JPIa	Asia	pig
AB094207	131JPIa	Asia	pig

AB094212	132JPIa	Asia	pig
AB094215	133JPIa	Asia	pig
AB094217	134JPIa	Asia	pig
AB094227	135JPIe	Asia	pig
AB094231	136JPIe	Asia	pig
AB094231	137JPIe	Asia	pig
AB094238	138JPIa	Asia	pig
AB094240	139JPIa	Asia	pig
AB094250	140JPIe	Asia	pig
AB094256	141JPIb	Asia	pig
AB094267	142JPIa	Asia	pig
AB094272	143JPIa	Asia	pig
AB094275	144JPIb	Asia	pig
AB094279	145JPIb	Asia	pig
AB094296	146JPIb	Asia	pig
AB094305	147JPIb	Asia	pig
AB094306	148JPIb	Asia	pig
AB094317	149JPIb	Asia	pig
AB096756	150JPIb	Asia	pig
AB105891	151JHSa	Asia	homo sapiens
AB105892	152JHSa	Asia	homo sapiens
AB105898	153JPIb	Asia	pig
AB105899	154JPIb	Asia	pig
AB105900	155JPIb	Asia	pig
AB105903	156JPIa	Asia	pig
AB105904	157JPIa	Asia	pig
AB107366	158JHSa	Asia	homo sapiens
AB107368	159JHSa	Asia	homo sapiens
AB112743	160JHSb	Asia	homo sapiens
AB115541	161JHSa	Asia	homo sapiens
AB115542	162JHSb	Asia	homo sapiens
AB115543	163JHSa	Asia	homo sapiens

AB115544	164JHSb	Asia	homo sapiens
AB154829	165JHSb	Asia	homo sapiens
AB154830	166JHSb	Asia	homo sapiens
AF195061	167SHSf	Europe	homo sapiens
AF195062	168SHSf	Europe	homo sapiens
AF195063	169SHSf	Europe	homo sapiens
AF296165	170TAPId	Asia	pig
AF296166	171TAPId	Asia	pig
AF296167	172TAPId	Asia	pig
AF332620	173NPIf	Europe	pig
AF336290	174NPIc	Europe	pig
AF336291	175NPIa	Europe	pig
AF336292	176NPIf	Europe	pig
AF336293	177NPIc	Europe	pig
AF336294	178NPIf	Europe	pig
AF336295	179NPIf	Europe	pig
AF336296	180NPIf	Europe	pig
AF336297	181NPIc	Europe	pig
AF336298	182NPIc	Europe	pig
AF336299	183NPIa	Europe	pig
AF466659	184UPIa	United States	pig
AF466660	185UPIa	United States	pig
AF466661	186UPIa	United States	pig
AF466662	187UPIa	United States	pig
AF466663	188UPIa	United States	pig
AF466664	189UPIa	United States	pig
AF466665	190UPIa	United States	pig
AF466667	191UPIa	United States	pig
AF466675	192UPIa	United States	pig
AF466676	193UPIa	United States	pig
AF466677	194UPIa	United States	pig
AF466678	195UPIa	United States	pig

AF466680	196UPIa	United States	pig
AF466681	197UPIa	United States	pig
AF466682	198UPIa	United States	pig
AF466683	199UPIa	United States	pig
AF466684	200UPIa	United States	pig
AF466685	201UPIa	United States	pig
AF503511	203UKPIe	Europe	pig
AF503512	204UKPIe	Europe	pig
AF516178	205KPIa	Asia	pig
AF516179	206KPIa	Asia	pig
AF527942	207KPIa	Asia	pig
AY032756	208NPIc	Europe	pig
AY032757	209NPIf	Europe	pig
AY032758	210NPIf	Europe	pig
AY032759	211NPIf	Europe	pig
AY323506	212SPIf	Europe	pig
AY362357	213UKHSe	Europe	homo sapiens
AY641398	214KHSa	Asia	homo sapiens
AY714267	215KHSa	Asia	homo sapiens
AY714268	216KHSa	Asia	homo sapiens
AY714269	217KHSa	Asia	homo sapiens
AY714270	218KHSa	Asia	homo sapiens
AY714271	219KHSa	Asia	homo sapiens
AY714272	220KHSa	Asia	homo sapiens

CHAPTER 3

Phylogeography and phylodynamics of european genotype 3 hepatitis e virus

Table 2S. Comparison between the tMRCAs and credibility intervals (95%HPD) of the main clades estimated by using the external (Nakano et al., 2012b) or the internal (heterochronous sequences) calibration approaches.

	External calibration						Internal calibration					
Node	**tMRCA**	**95%HPDL**	**95%HPDU**	**Year**	**95%HPDU**	**95%HPDL**	**tMRCA**	**95%L**	**95%U**	**Year**	**95%HPDU**	**95%HPDL**
Root	199	116	289	1810	1721	1894	186	46	527	1822	1482	1963
A	120	71	172	1889	1838	1939	113	28	331	1895	1678	1981
HEV-3f	74	40	192	1935	1818	1970	58	16	199	1950	1810	1993
HEV-3e	92	54	129	1917	1881	1956	88	22	263	1920	1746	1987
B	114	70	157	1895	1853	1940	107	29	308	1901	1701	1980
HEV-3c	90	54	125	1919	1885	1956	85	23	250	1923	1759	1986
B'	90	56	124	1920	1886	1954	84	21	243	1924	1766	1988
HEV-3a	50	30	66	1959	1944	1980	45	13	129	1963	1880	1996
HEV-3d	7	0,8	9,3	2002	2000	2009	4	0,8	12	2005	1997	2008
HEV-3b	66			1944			52	14	178	1956	1831	1995

SCIENTIFIC CONTRIBUTIONS

Publications

- Luzzago C, Lauzi S, Ebranati E, Giammarioli M, Moreno A, Cannella V, MasoeroL, Canelli E, Guercio A, Caruso C, Ciccozzi M, De Mia GM, Acutis PL, Zehender G, Peletto S. 2014 Extended genetic diversity of bovine viral diarrhea virus and frequency of genotypes and subtypes in cattle in Italy between 1995 and 2013.Biomed Res Int. 2014:147145. doi: 10.1155/2014/147145.

- Zehender G, Ebranati E, Lai A, Luzzago C, Paladini S, Tagliacarne C, Galli C, Galli M, Ciccozzi M, Zanetti AR, Romanò L. 2014 Phylogeography and phylodynamics of European genotype 3 hepatitis E virus. Infect Genet Evol. 25:138-43.

- Zehender G., Ebranati E., Shkjezi R., Papa A., Luzzago C., Gabanelli E., Lo Presti A., Lai A., Rezza G., Galli M., Bino S., Ciccozzi M. 2013 Bayesian Phylogeography of Crimean-CongoHemorrhagic Fever Virus in Europe PLoS ONE 8(11): e79663.

International Meeting

- **Annual meeting of the Society for Molecular Biology and Evolution – Vienna, Austria, 12 - 16 luglio 2015**. (Poster: Cerutti F, Ebranati E, Luzzago C, Lauzi S, Caruso C, Masoero L, Giammarioli M, Moreno A, Cannella V, Canelli E, Guercio A, Ciccozzi M, De Mia GM, Acutis PL, Zehender G,Peletto S. Evolutionary Dynamics of Bovine Viral Diarrhoea Virus 1 (BVDV-1) in Italy: phylogeography and contact networks)

- ***Joint U.S. BVDV/ESVV Pestivirus Symposium Pestiviruses: Old Enemies, New Challenges,Oct. 14 and 15 2014, in Kansas City, MO,*** *abstract book, 6. (Oral presentation: Luzzago C, Lauzi S, Ebranati E, Giammarioli M, Moreno A, Cannella V, MasoeroL, Canelli E, Guercio A, Caruso C, Ciccozzi M, Petrini S, De Mia GM, Acutis PL, Zehender G, Peletto S. Is Italy one of the countries with the highest genetic diversity of bovine viral diarrhea virus?)*

- **Chamois International Congress, Parco Nazionale della Maiella, Lama dei Peligni 17-20 giugno 2014** (poster: Spatial and temporal phylogeny of Border disease virus in Pyrenean chamois Luzzago C., Ebranati E., Lanfranchi P., Cabezón O., Lavín S., Rosell R.,Rossi L., Zehender G., Marco I.)

- **II International Rupicapra Symposium- Biology, Health, Monitoring and Management- Bellver de Cerdanya (Catalonia, Spain) 24-25/10/13** (poster: Virulence-associated genes of *Staphylococcus aureus* isolates from soft tissue infections in Alpine chamois C. Luzzago, C. Locatelli, L. Scaccabarozzi, R. Viganò, G. Sironi, M. Besozzi, N. Formenti, B. Castiglioni, P. Lanfranchi and P.Cremonesi)

National Meeting

- ***XVI Congresso Nazionale S.I.Di.L.V. -Montesilvano (PE), 30 Settembre - 2 Ottobre 2015*** *(oral presentation: Cerutti F, Ebranati E, Luzzago C, Lauzi S, Caruso C, Masoero L, Giammarioli M, Moreno A,Cannella V, Canelli E, Guercio A, Ciccozzi M, De Mia GM, Acutis PL, Zehender G,Peletto S. Filodinamica, filogeografia e caratterizzazione molecolare full genome di Bovine viral diarrhoea virus (BVDV))*

- **V Workshop Nazionale di Virologia Veterinaria, Teramo, 26-27 giugno 2014** (oral presentation: *Luzzago C, Lauzi S, Ebranati E, Giammarioli M, Moreno A, Cannella V, MasoeroL, Canelli E, Guercio A, Caruso C, Ciccozzi M, De Mia GM, Acutis PL, Zehender G, Peletto S.Variabilità genetica di BVDV e frequenza di genotipi e sottotipi in Italia nel periodo 1995-2013)*

ACKNOWLEDGEMENTS

This research work would not have been possible without the contribution of a great and effective interdisciplinary team of competence.

I would like to thank all of them and in particular:

- Gianguglielmo Zehender and Erika Ebranati of the Department of Biochemical and Clinical Sciences “Luigi Sacco”, Section of Infectious Diseases, University of Milan;
- Valerio Bronzo of the Department of Health, Animal Science and Food Safety, University of Milan;
- Stefania Lauzi of the Department of Veterinary Science and Public Health, University of Milan;
- Ignasi Marco Sánchez and all the research group of the Servei d'Ecopatologia de Fauna Salvatge, Facultat de Veterinària Campus de la UAB · Barcelona · Spain;
- Simone Peletto and Pier Luigi Acutis, Francesco Cerutti, Loretta Masoero, Claudio Caruso of the Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d’Aosta;
- Gian Mario De Mia and Monica Giammarioli of the “Centro di Referenza Nazionale per lo Studio delle Malattie da Pestivirus e Asfivirus”, Perugia;
- Ana Moreno of the “Istituto Zooprofilattico Sperimentale della Lombardia e dell’Emilia-Romagna;
- Vincenza Canella of the Istituto Zooprofilattico Sperimentale della Sicilia.

My research activity has been supported by:

- Italian Ministry of University and Research in the framework of the research project “Genomics and host-pathogen interactions: a model study in the One-Health perspective (PRIN project n. 2010P7LFW4)

- Italian Ministry of Health in the framework of the Research Project “Phylodynamic, phylogeography and full genome characterization of bovine viral diarrhea virus” (Project IZS PLV 16/11RC).

I would like to thank again my co-tutor, Gianni Zehender, for his curiosity and knowledge of evolution of infectious diseases in all the host species (chamois included) and also Erika Ebranati for the invaluable collaboration on phylogenetic analyses. My experience with their research group has been successful and I hope that will continue in the future. Thanks again also to my colleague and friend Stefania Lauzi for all the time spent together with me, trees and sequences.

Thanks to my wonderful family, Attilio, Ernesto, Augusto and Paolo. They have supported, and tolerated a crazy mamma and sposa.

Be a virus see the world.