



journal homepage: [www.elsevier.com/locate/csbj](http://www.elsevier.com/locate/csbj)

# Mode and tempo of human hepatitis virus evolution

Rachele Cagliani<sup>1</sup>, Diego Forni<sup>1</sup>, Manuela Sironi<sup>\*</sup>

Bioinformatics, Scientific Institute, IRCCS E. MEDEA, 23842 Bosisio Parini, Lecco, Italy



## ARTICLE INFO

### Article history:

Received 29 July 2019

Received in revised form 19 September 2019

Accepted 21 September 2019

Available online 25 October 2019

### Keywords:

Human hepatitis virus

Zoonosis

Molecular dating

Host switch

## ABSTRACT

Human viral hepatitis, a major cause of morbidity and mortality worldwide, is caused by highly diverse viruses with different genetic, ecological, and pathogenetic features. Technological advances that allow throughput sequencing of viral genomes, as well as the development of computational tools to analyze such genome data, have largely expanded our knowledge on the host range and evolutionary history of human hepatitis viruses. Thus, with the exclusion of hepatitis D virus, close or distant relatives of these human pathogens were identified in a number of domestic and wild mammals. Also, sequences of human viral strains isolated from different geographic locations and over different time-spans have allowed the application of phylogeographic and molecular dating approaches to large viral phylogenies. In this review, we summarize the most recent insights into our understanding of the evolutionary events and ecological contexts that determined the origin and spread of human hepatitis viruses.

© 2019 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Contents

1. General introduction .....	1384
2. Introduction to molecular dating .....	1385
3. Hepatitis A virus .....	1385
4. Hepatitis B virus .....	1386
5. Hepatitis C virus .....	1387
6. Hepatitis D virus .....	1390
7. Hepatitis E virus .....	1390
8. Conclusions .....	1392
Declaration of Competing Interest .....	1393
Acknowledgement .....	1393
References .....	1393

## 1. General introduction

The worldwide burden of viral hepatitis in terms of death and disability is enormous. In 2015, viral hepatitis caused 1.34 million deaths, the majority of which imputable to the effects of chronic

HBV (hepatitis B virus) and HCV (hepatitis C virus) infection [1]. An estimated 5% of HBV-infected persons are also co-infected with HDV (hepatitis delta virus), which worsens the clinical outcome compared to HBV monoinfection [1]. Less than 5% of overall hepatitis mortality is caused by HAV (hepatitis A virus) and HEV (hepatitis E virus), that are usually associated with acute, self-limiting disease [1]. However, the case fatality rate of HEV is particularly high in specific groups, notably pregnant women and elderly individuals [2]. Although rare, infection with HAV can also cause acute liver failure and death, and the risk increases with age [3]. Despite the existence of a safe and effective vaccine, HAV remains a common cause of acute viral hepatitis in many regions of the world [4].

Abbreviations: ORF, open reading frame; NHP, non-human primates; TDRP, time-dependent rate phenomenon; STI, sexually transmitted infection; TMRCA, time to the most recent common ancestor; RdRp, RNA-dependent RNA polymerase.

\* Corresponding author.

E-mail address: [manuela.sironi@bp.inf.it](mailto:manuela.sironi@bp.inf.it) (M. Sironi).

<sup>1</sup> These authors equally contributed.

**Table 1**  
Human hepatitis viruses: taxonomic classification and general features.

	Family	Genus	Genome type	Genome length (kb)	Natural host	Course of infection
HAV	<i>Picornaviridae</i>	<i>Hepatovirus</i>	Linear ssRNA(+)	7.5	Human	Acute
HBV	<i>Hepadnaviridae</i>	<i>Orthohepadnavirus</i>	Partially circular dsDNA	3.2	Human, non-human primates	Acute/chronic
HCV	<i>Flaviridae</i>	<i>Hepacivirus</i>	Linear ssRNA(+)	9.6	Human	Acute/chronic
HDV	Unassigned	<i>Deltavirus</i>	Circular ssRNA(-)	1.7	Human	Acute/chronic
HEV	<i>Hepeviridae</i>	<i>Orthohepevirus</i>	Linear ssRNA(+)	7.2	Human, rabbit, pig, wild boar, camel, deer, mongoose, and other mammals	Acute

Human hepatitis viruses are extremely diverse and consequently belong to different viral families (Table 1). In recent years, the availability of high-throughput technologies has revealed that relatives of human hepatitis viruses can be found in a wide variety of animals. This finding, as well as the increasing availability of the genome sequences of human-infecting viruses sampled across different geographic areas, has largely expanded our knowledge about the genetic diversity and evolutionary origin of these human pathogens. In this Review, we thus focus on the latest insights into the possible events and ecological contexts that determined the origin and spread of human hepatitis viruses. A short presentation of the most widely used approaches to estimate the ages of viral lineages is also provided to contextualize recent research efforts on these viruses.

## 2. Introduction to molecular dating

Molecular dating analyses using virus genetic data can be particularly informative due to the rapid rate of evolution of many viral species. By converting genetic differences among sequences into time units, molecular dating provides information on the timing of viral spread or emergence. Most molecular dating approaches are based on maximum-likelihood or Bayesian phylogenetic frameworks [5,6], and they usually exploit two strategies: molecular clock calibration using the sampling dates of the viral sequences (tip dating) and/or calibration using information on some internal nodes of the phylogeny. Tip dating is well-suited to study relatively recent events (e.g., epidemics or intra-host evolution), but requires that a temporal signal is present in the dataset (i.e., that the sequences have accumulated a measurable amount of change between sampling times) [6]. Calibration using internal nodes can in principle allow to dig deeper into the past, but requires some *a priori* knowledge about the virus evolutionary history (e.g., host-virus co-evolution, paleontological information). The widespread use of molecular dating has however revealed that the relationship between genetic divergence and time is complex, as evolutionary rates tend to vary with the time frame of measurement. In particular, high evolutionary rates are observed in the short term, whereas low rates are inferred in long time span studies [7–9]. This pattern was observed for many viral lineages and is sometimes referred to as the time-dependent rate phenomenon (TDRP) [10]. Failure to account for the TDRP can potentially lead to erroneous molecular dating results [10,11].

The TDRP reflects mutation rate in very short timescales and substitution rate in very long timescales, during which transient deleterious mutations are removed by the action of natural selection, leading to lower rate estimates [12]. Another factor most likely contributing to the TDRP is the saturation of nucleotide substitutions, which is extremely rapid in viral genomes, especially when the polymerase is error-prone [12]. Thus, recent analyses indicated that, for all Baltimore classification groups, viral evolutionary rates tend to decrease continuously with the timescale of measurement [10,11]. Because the rate of decay is consistent with a power law relationship between substitution rate and sampling

timescale, a model using a simple regression was at first proposed to estimate the TDRP effect on viral phylogenies [10]. Very recently, this approach was implemented in a Bayesian statistical framework, in which evolutionary rates can vary among different time epochs [13]. Before the introduction of such an approach, effective attempts to correct for the TDRP were performed by the use of nucleotide substitution models that allow site- and branch-specific variation in selective pressure (selection-informed models). These models, which were applied to analyze the ancient evolution of some viral lineages, at least partially correct for the effects of both purifying selection and substitution saturation in branch length estimation [14–16].

## 3. Hepatitis A virus

HAV is mainly transmitted via the faecal-oral route through exposure to contaminated food or water, or through direct contact with infected people. HAV is a single-strand, positive RNA virus with a genome of approximately 7.5 kb in length (Table 1). The HAV genome contains a single ORF flanked by a relatively long 5' UTR and a 3' UTR. The 5' UTR harbors an internal ribosome entry site that directs the cap-independent translation of HAV proteins. The ORF encodes a polyprotein processed in 11 mature proteins: 5 structural proteins involved in capsid formation (VP4, VP2, VP3, VP1, and pX, deriving from P1 segment) and 6 non-structural proteins with a role in RNA genome amplification (2B, 2C, 3A, 3B, 3Cpro, and 3Dpol, deriving from P2 and P3 segments) [4,17].

Based on genomic structure, HAV belongs to the family *Picornaviridae* within the genus *Hepatovirus*. Nevertheless, many characteristics distinguish HAV (and hepatoviruses in general) from other *Picornaviridae* family members. Some peculiar features include the primary tropism for hepatocytes, the ability to shed as nonenveloped virus in feces and as enveloped particles in blood, as well as some genomic features such as low G/C ratio, low CpG levels, and strong codon bias [18,19].

HAV was identified as the etiologic agent of hepatitis A by Feinstone and colleagues [20] in 1973. Unlike HBV and HCV, which establish chronic infections in humans, HAV infection is usually acute and generates lifelong immunity. This condition is able to determine the disappearance of the virus in small and isolated populations [21,22] and did not probably favor its maintenance in early human communities. It is thus legitimate to wonder how HAV survived and evolved during early human history, a question that remains presently unanswered.

For a long time, it was thought that hepatoviruses were restricted to humans and non-human primates (NHPS), with genetically distinct variants classified as six main different genotypes [23]: three isolated from humans (HAV, genotype I–III) and sub-classified in 6 subgenotypes (IA, IB, IIA, IIB, IIIA, IIIB) and three of simian origin (SHAV, genotype IV–VI). However, despite genetic heterogeneity, HAV viruses belong to a single common serotype.

In recent years, the advent of new sequencing approaches has led to an exponential increase in the identification of new viral spe-

cies, including highly diverse non-primate hepatoviruses. Several HAV-related viruses were identified in different mammalian orders. In particular, a number of HAV-like viruses were recovered in placental mammals, mainly in bats and rodents, but also in tree shrews, hedgehogs, seals and Chinese woodchucks [24–26]. Recently, de Oliveira Carneiro and colleagues [27] identified a novel HAV-related virus in *Didelphis aurita*, a Brazilian common opossum, further extending the host range of mammal-infecting hepatoviruses. Moreover, viruses related to mammalian hepatoviruses were detected in reptiles, amphibians, and fish [28]. These advances allowed new insights into the evolutionary history of HAV.

The phylogenetic relationships among hepatovirus that infect small mammals only partially reflects those among their hosts, suggesting multiple, non-recent cross-species and cross-order host switches during hepatovirus evolution [24]. This observation is supported by recombination events observed in hepatoviruses that have been identified in genetically and geographically distant hosts [29]. These cross-species transmission events also involved the opossum hepatovirus, which most likely originated from an ancestral host switch from rodents into marsupials [27]. Conversely, hepatovirus phylogenies suggest no host switch involving a primate donor. This evidence, the absence of recombination events between HAVs and non-primate hepatoviruses, as well as the observation that primate hepatoviruses form, regardless of the genomic region considered, a monophyletic group in the topology of hepatovirus phylogenies, support the hypothesis that humans and NHP have acquired hepatoviruses from other animal reservoirs relatively recently [24,29]. However, if and when this hypothetical host-jump occurred into primates remains to be clarified.

Phylogenetic and ancestral state reconstruction suggested a likely cricetid rodent origin for primate HAVs and marsupial hepatoviruses, whereas a laurasiatherian host origin was proposed for all mammalian hepatoviruses [24,27] (Fig. 1). In this scenario, the evolutionary history of hepatoviruses is evocative of that of hantaviruses, as the origin of mammalian hantaviruses is traced back to bats and insectivores [30]. Thus, the supposed origin of hepatoviruses in insectivorous laurasiatherian mammals, as well as the preservation of some structural and functional characteristics similar to present-day insect picorna-like viruses (*Dicistroviridae*) [31] led Drexler and colleagues to hypothesize a more ancient evolutionary origin of HAVs, with an ancestry in a primordial insect-borne virus [24].

In conclusion, HAV emergence in humans likely represents a relatively recent evolutionary event, probably of zoonotic origin. Nonetheless, the ancestor of human hepatoviruses has yet to be identified. The characterization of other hepatoviruses in primates, and mammals in general, will be instrumental to the identification of the HAV ancestors and will clarify the evolutionary history of hepatoviruses.

#### 4. Hepatitis B virus

HBV was the first human hepatitis virus to be isolated and identified in 1970 [32]. HBV transmission varies depending on the prevalence of infection. In areas with a low prevalence (<2%), the most common mode of transmission is through infected blood or high-risk behaviors (e.g. unprotected sex or injecting drug use). In high- and intermediate- prevalence areas, HBV is commonly spread through perinatal and horizontal (especially among children) routes [33]. HBV belongs to the *Hepadnaviridae* family, which comprises two genera: *Orthohepadnavirus* (mammal-infecting viruses) and *Avihepadnavirus* (bird-infecting viruses) (Fig. 2A). Its genome, a partially double-stranded circular DNA of about 3.2 kb (Table 1), is composed of four overlapping frameshifted open

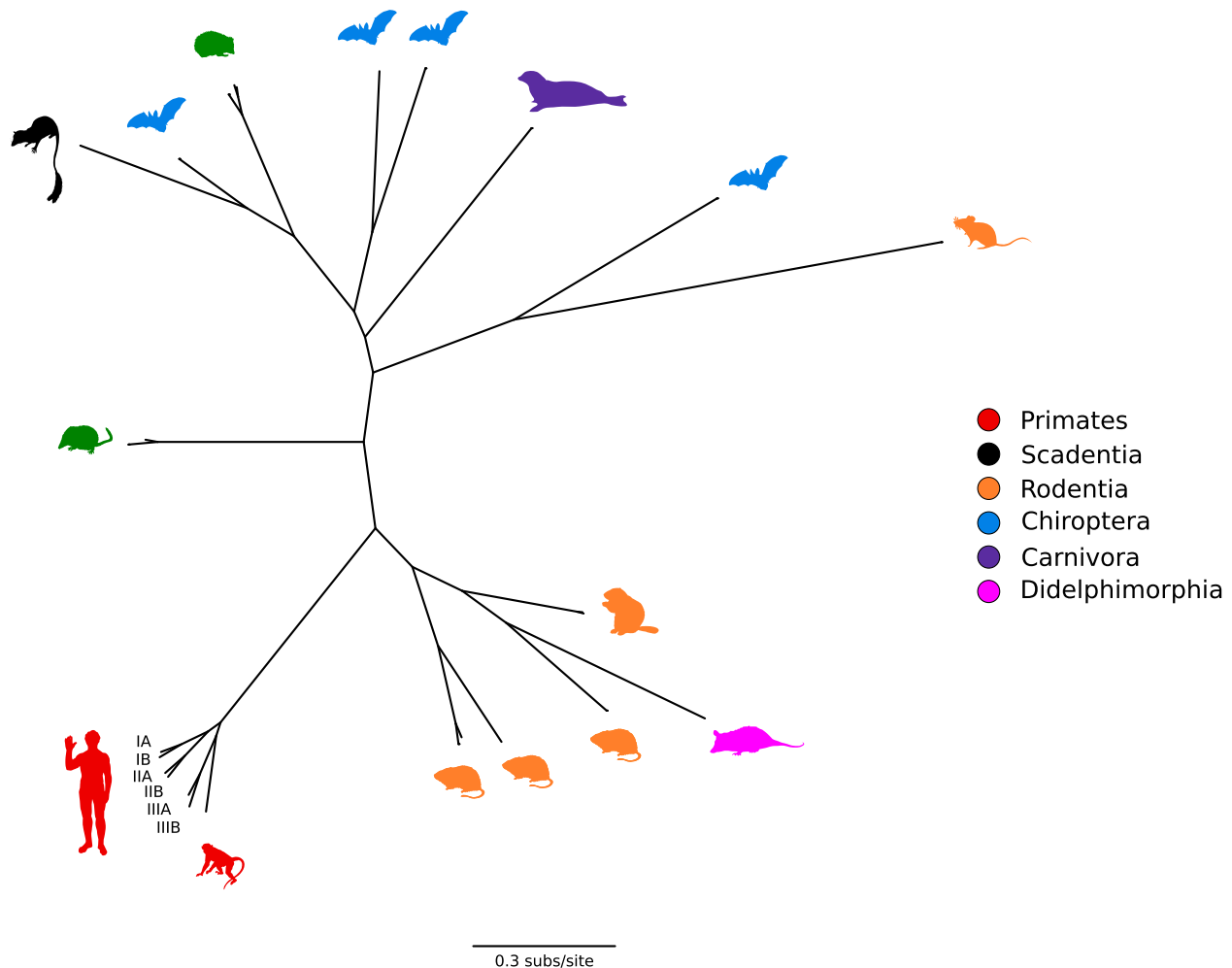
reading frames (ORFs) [34]. Viral replication is carried out by a reverse transcriptase with no proofreading ability and considerable variability exists among strains. Thus, at least nine genotypes (A–I) plus a tentative one (J), with a heterogeneous global distribution, have been described to date [35–39] (Fig. 2B).

Despite its worldwide diffusion and the accumulation of detailed knowledge on the associated pathologies, the origin and evolutionary history of HBV are still debated [34,40]. Hepadnaviruses were detected in several mammals, including NHPs, rodents and bats, birds, fish, and reptiles [41,42] (Fig. 2A). Recently, Lauber et al. discovered a family of fish viruses with genomic features similar to those of HBV, dating the origin of the *Hepadnaviridae* family to at least 300 million years ago [43]; this finding, together with the discovery of endogenous hepadnavirus elements integrated in the genome of birds and reptiles [44–47], suggests a long and complex relationship between this viral family and its hosts.

Concerning HBV, different theories were proposed to explain its origin, but all of them have some sort of limitation. HBV was initially thought to have emerged quite recently in the New World from genotypes F/H infecting Amerindians [48,49] (Fig. 2A and B). However, the discovery of an ancient strain in a 16th century Asian mummy, as well as the worldwide diffusion of hepadnaviruses in NHPs, questioned this hypothesis [50] (Fig. 2A). An alternative theory suggests that HBV followed the Out-of-Africa migration of modern humans, which occurred approximately 60,000 years ago [51–53]. In particular, Paraskevis et al. found a good correspondence between the demographic histories of HBV and those of human populations [53]. These authors also showed that the substantial divergence of the F and H genotypes (Fig. 2A), a major evidence in favor of the New World origin hypothesis [49], was probably due to positive selection acting on those branches [54]. Nonetheless, the extensive application of ancient DNA sequencing revealed a more complex scenario. In fact, two European Neolithic HBV genomes did not cluster with any extant human strain in the phylogenetic tree, but did cluster with NHP viruses [55] (Fig. 2A). Other authors [9], who sequenced 12 ancient HBV genomes of different ages (800–4500 years old), obtained a similar result, with the ancient strains clustering with known modern genotypes or forming new clades (Fig. 2A). This implies that some HBV lineages of the past went extinct (Fig. 2A). Moreover, Muhlemann and coworkers showed that the geographic distribution of ancient samples does not match the modern genotype distribution [9]. They thus suggested that early evolutionary scenarios can be concealed and overwritten by more recent migratory events [9].

A third hypothesis for the origin of HBV posits that hepadnaviruses co-specified with their primate hosts in the New World and in the Old World. Thus, multiple zoonotic transmissions would have originated HBV genotypes found in humans [56]. This scenario is supported by the diffusion of hepadnaviruses in diverse primate species and by the inferred divergence time of the *Orthohepadnavirus* and *Avihepadnavirus* genera, that is very similar to that of their host classes [56]. However, the recent identification of a novel hepadnavirus in capuchin monkeys confirmed that New World monkeys are infected by viruses that are very distantly related to HBV (Fig. 2A), indicating that they do not represent the direct ancestors of genotypes H and F [41]. Instead, evolutionary analyses with human and NHP viral strains placed the origin of HBV ancestors in Hominoid Old World primates, preceding the formation of the human lineage [41].

In summary, although considerable progress was achieved in recent years, a high level of uncertainty concerning the ultimate origin of HBV still exists. The particular organization of the viral genome (i.e. overlapping reading frames in a short genome) limits the variability of most of nucleotide positions (i.e. to avoid the introduction of nonsynonymous mutations) and results in a



**Fig. 1.** Hepatovirus RdRp domain phylogeny. A maximum-likelihood phylogenetic tree of representative hepatoviruses [24,27] was generated using PhyML [136]. Hepatovirus host silhouettes are colored according to taxonomic order. The human HAV subgenotypes are also reported.

relatively slow mutation rate. This characteristic, along with the action of natural selection on particular genotypes [54], as well as the adaptation to different human populations [34], contributes to HBV variability and complicates inferences about its origin. Finally, different studies [9,50,57] have shown that, as for other viruses (see Section 2), HBV substitution rates are affected by viral sampling time frames. Indeed, the evolutionary rates generated using information from ancient HBV genomes were shown to fit well with the TDRP regression line calculated for Baltimore Group VI and VII viruses (i.e., reverse-transcribing viruses) [11]. Crucially, these results indicate that, whereas tip calibration approaches have demonstrated to be useful in the reconstruction of recent epidemiological events [58–60], limiting analysis to extant strains for the reconstruction of ancient HBV evolution can be misleading [9] and that approaches that correct for the TDRP should be envisaged.

## 5. Hepatitis C virus

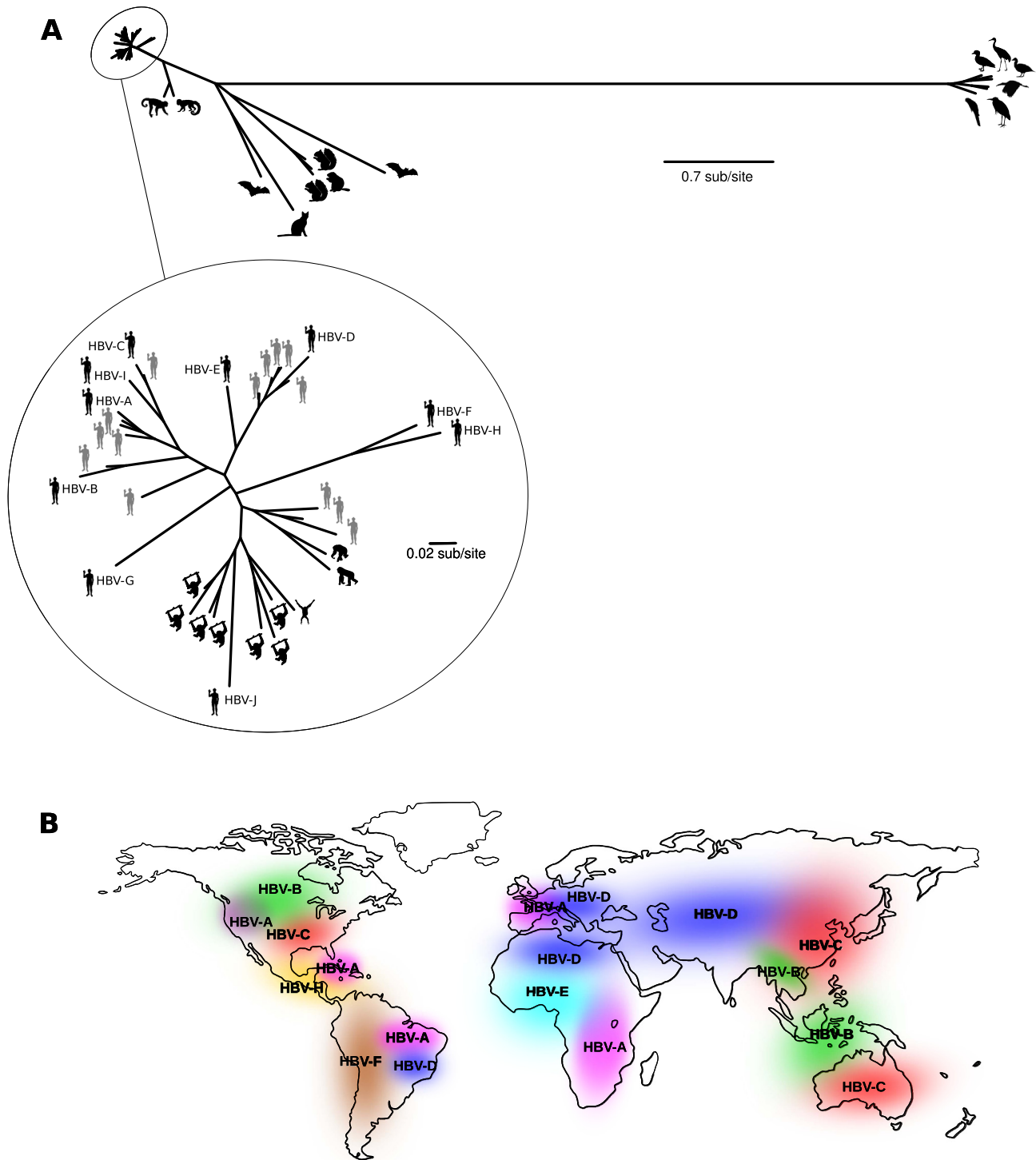
HCV is an enveloped virus belonging to the *Flaviviridae* family (genus *Hepacivirus*). In analogy to other members of the family, HCV has a 9.6 kb positive-stranded linear RNA genome. The virus encodes a single polyprotein that is processed by cellular and viral proteases to yield at least 10 mature products.

HCV was identified in 1989 by Houghton and colleagues as a cause of non-A and non-B hepatitis [61]. If left untreated, HCV

can persist lifelong in humans, often resulting in cirrhosis and hepatocellular carcinoma.

Presently, the HCV worldwide seroprevalence is estimated to be ~1% [1], with about 71 million persons living with chronic infection [1]. The HCV epidemic apparently started recently, in the 1930s–1940s, as a consequence of practices that determined parenteral or percutaneous exposure (e.g., blood transfusion, vaccination campaigns, and intravenous drug injection) [62–64]. For instance, one of the most affected countries is Egypt, where the virus was most likely disseminated through nationwide vaccination programs or contaminated blood-derived products [65]. In fact, sexual or vertical transmission of HCV are relatively rare, and the overwhelming majority of infections occur via the parenteral/percutaneous route. Thus, due to historical reasons and to the transmission pattern, a small number of so-called “epidemic” HCV subtypes (1a, 1b, 3a, and 2a) account for most infections worldwide [64,66].

HCV is, however, genetically heterogeneous and the epidemic subtypes represent a minor fraction of viral diversity. Eight major HCV genotypes (HCV-1 to -8) have been described, and these are further divided into at least 90 subtypes ([https://talk.ictvonline.org/ictv\\_wikis/flaviviridae/w/sg\\_flavi/56/hcv-classification](https://talk.ictvonline.org/ictv_wikis/flaviviridae/w/sg_flavi/56/hcv-classification)) (Fig. 3A). Several of these genotypes and subtypes were identified and classified only recently [67,68], suggesting that a considerable proportion of HCV diversity remains undescribed. Moreover, a number of natural inter-genotype recombinants were reported

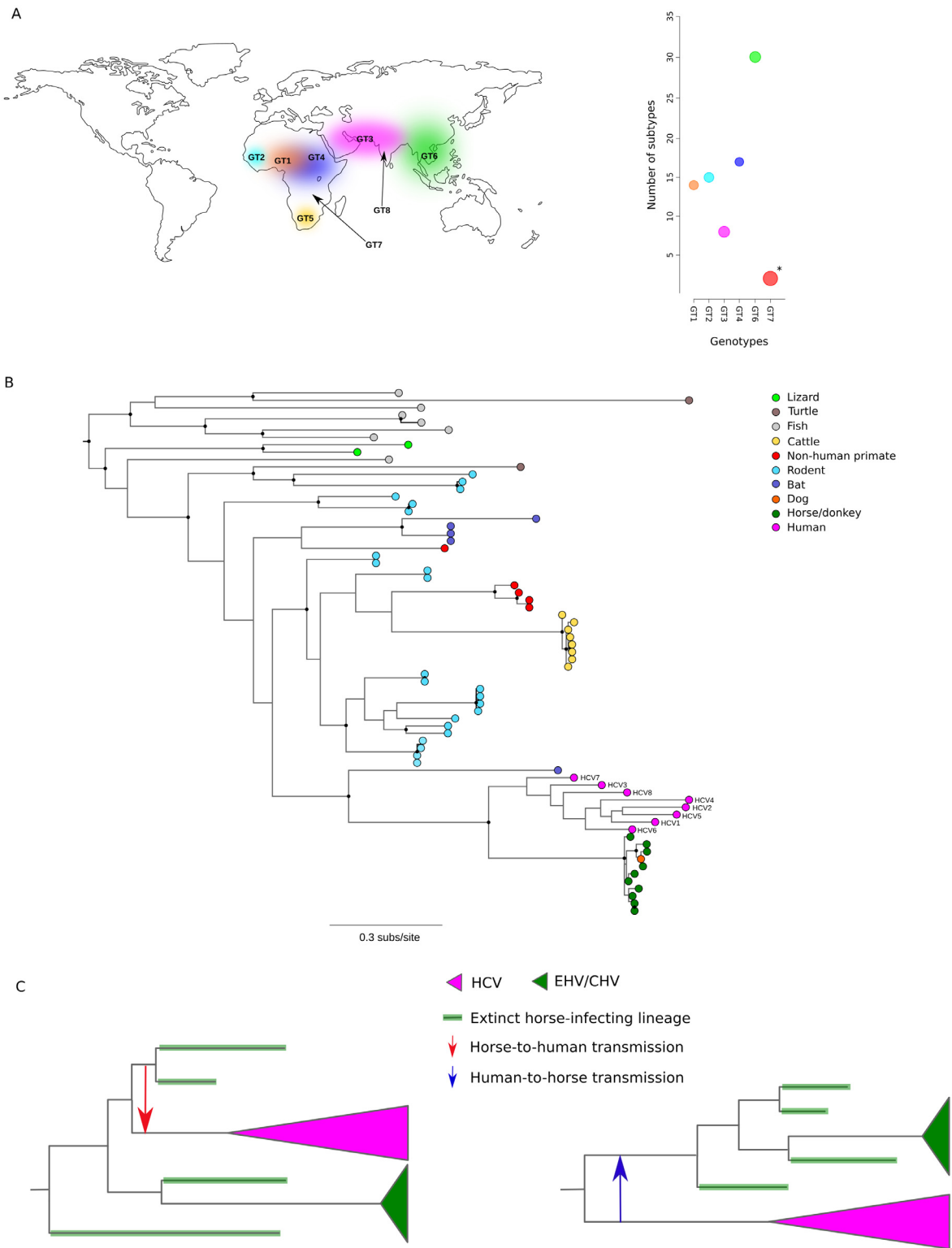


**Fig. 2.** Hepadnavirus diversity and geographic distribution. (A) Maximum-likelihood phylogenetic tree of known hepadnaviruses. Symbols are indicative of viral hosts. The tree of representative hepadnavirus polymerase/reverse-transcriptase (RT) domain was obtained using PhyML [136]. The insert shows the phylogenetic relationship of HBV extant genotypes, HBV ancient strains (in gray), and NHP viruses. (B) Present-day worldwide distribution of the most prevalent HBV genotypes. Geographic distributions derive from Refs. [35–39,137].

([https://talk.ictvonline.org/ictv\\_wikis/flaviviridae/w/sg\\_flavi/38/table-4-recombinant-rf-hcv-genomes](https://talk.ictvonline.org/ictv_wikis/flaviviridae/w/sg_flavi/38/table-4-recombinant-rf-hcv-genomes)).

HCV genotypes display antigenic variability and viral genetic diversity is geographically structured: in sub-Saharan Africa and South-East Asia highly divergent subtypes of the same genotype dominate transmissions across contiguous areas (Fig. 3A). These subtypes are referred to as “endemic” [62,64] and their presence

is consistent with a long-standing association of HCV with populations living in these regions. Because parenteral exposure became common only in the relatively recent past, several hypotheses were formulated to account for the maintenance of endemic HCV transmission. Some authors proposed that traditional practices such as circumcision, tattooing, piercing or acupuncture facilitated and maintained HCV infection among human populations [64,69].



**Fig. 3.** HCV genetic diversity and relationships with other hepaciviruses. (A) Worldmap showing the areas where endemic transmission of different genotypes was reported (data from Ref. [138]). Graphical representation of genotype (GT) diversity. The number of distinct recognized subtypes is represented on the Y axis. Circle size is proportional to the average pairwise distance between subtypes. Genotype 7 is marked with an asterisk as only two subtypes are known. (B) Phylogenetic tree of the RdRp domain of known hepaciviruses. The tree was obtained using RaxML with 1000 bootstrap replicates [139]. Nodes with support equal to or higher than 0.7 are marked with a black dot. (C) Hypothetical viral phylogenies that illustrate the effect of viral lineage extinction on the evolutionary inference about the origin of HCV and EHV/CHV. In the left panel, a horse-to human transmission event is hypothesized, with the following extinction of the transmitted EHV lineage. In the right panel, a reverse zoonosis introduces a HCV-like virus in horse populations; the following extinction of several EHV lineages accounts for the low genetic diversity of extant strains.

Others indicated that sexually transmitted infections (STIs) that disrupt mucosal integrity are responsible for increased sexual HCV transmission [70]. This was indeed shown to be the case in

modern high-risk populations [71] and STIs have probably been common throughout human history [72]. An alternative scenario is that the bite of arthropods, especially those taking large blood

meals, can mechanically transmit HCV, possibly from other animal reservoirs such as horses [73,74]. HCV infection is in fact restricted to our species but, thanks to extensive field work, a number of hepaciviruses have been described in domestic and wild mammals, as well as in reptiles and fish [28,75–84] (Fig. 3B). At present, the largest diversity of hepacivirus species seems to be hosted by rodents and bats (Fig. 3B). Instead, as previously noted [75], the lowest genetic diversity is observed for hepaciviruses that infect cattle and horses (Fig. 3B), suggesting that husbandry practices may have resulted in the artificial selection of specific viral strains or facilitated recent viral transmission from some other animal source (e.g., commensal rodents). Non-human primates also host hepaciviruses, but these are highly divergent from HCV. Overall, the phylogenetic relationships among hepaciviruses poorly mirror those among their hosts (Fig. 3B), suggesting several cross-species and cross-order host switches during viral evolution. Up to now, the closest relatives of HCV were identified in horses/donkeys (equine hepacivirus, EHV) and dogs (canine hepacivirus, CHV) (Fig. 3B). Because CHV is less genetically diverse than EHV, the canine virus possibility originated as a recent cross-species transmission from horses [85], hinting at the ability of hepacivirus to shift among genetically distant hosts. Despite these advances, the events that led to the origin of HCV are still unknown. Taking as a fact the relatedness of the human virus to EHV/CHV, possible scenarios include that: i) HCV originated from a cross-species transmission of EHV, ii) that EHV was transmitted to horses by humans infected with HCV, which leaves the question on HCV origin open; iii) that HCV and EHV originated from the cross-species transmissions of the same (or similar) virus, with subsequent host adaptation and divergence. If this were the case, multiple cross-species transmission events may have originated distinct HCV genotypes [85].

Teasing apart these possibilities clearly requires understanding of the timing and circumstances of HCV (and EHV) evolution. Up to date, no archaeological sample carrying traces of HCV or EHV has been described and the oldest HCV sequence dates to 1953 [86] (1979 for EHV [82]). Thus, molecular dating efforts have relied on extant sequences, with the difficulties associated with the TDRP. Studies that did not account for the TDRP provided estimates of the time to the most recent common ancestor (TMRCA) of HCV genotypes in a range between ~200 and 1000 years ago [63,64,76,87,88]; the origin of the horse virus was dated around 1800 CE [85]. A study that separately accounted for the rate of synonymous and nonsynonymous substitutions estimated HCV to have originated at least 2000 years ago [89]. Recently, a method based on an a selection-informed model was used to estimate the divergence time of HCV genotypes and the origin of extant EHV/CHV strains [70]. This approach, provided estimates of ~3000 years ago for the TMRCA of extant HCV genotypes (with a low-bound estimate of ~5000 years before present) and of ~800 years ago for EHV/CHV [70]. If these dates are taken to provide at least an indication of the real evolutionary scenario, the possibility that HCV was transmitted to humans by horses infected with EHV can be excluded, an observation in line with the low diversity of EHV/CHV [82]. The origin of EHV/CHV as reverse zoonosis (i.e., the transmission of a human virus to animals) seems also unlikely, as in this case horse viruses should cluster within HCV diversity, unless the HCV lineage that originated EHV went extinct in the last 800 years. Indeed, as the HBV story exemplifies, viral lineage extinction can occur and was previously documented for other human pathogens such as parvovirus B19 [90], and variola virus [91]. As anticipated above, breeding practices may facilitate this process in the case of animal viruses. We know, for instance, that a minimum of two horse lineages went extinct during the domestication process and that horse genetic diversity has largely been shaped by events that occurred in the last few

centuries [92]. It is thus possible that human-mediated selection on the host also resulted in the artificial selection of viral lineages. This would explain the relatively recent origin of extant EHV strains and their low diversity. If viral lineage extinction did occur, the time frames of EHV and HCV evolution would be underestimated and the scenario of HCV originating as a zoonosis from horses (or EHV as a reverse zoonosis) may still hold (Fig. 3C). Of course, the alternative possibility that EHV and HCV were transmitted independently to their present-day hosts by a third unknown reservoir is also in line with data on extant diversity and, if the transmission to horses occurred recently, does not require to postulate viral lineage extinction. Thus, a number of open questions remain on the origin of HCV. Hopefully, technological advances that allow sequencing of trace genetic material from humans and horses back in the past. At the same time, the extensive application of metagenomic approaches to different animal hosts across diverse geographic regions will expand our knowledge on hepacivirus diversity and eventually uncover the direct ancestor of HCV (if it still exists). Indeed, the possibility that the different HCV genotypes derived from independent cross-species transmission events [85] would imply that viruses related to HCV are relatively common in the wild, thus giving good chances to be recovered in large field surveys.

## 6. Hepatitis D virus

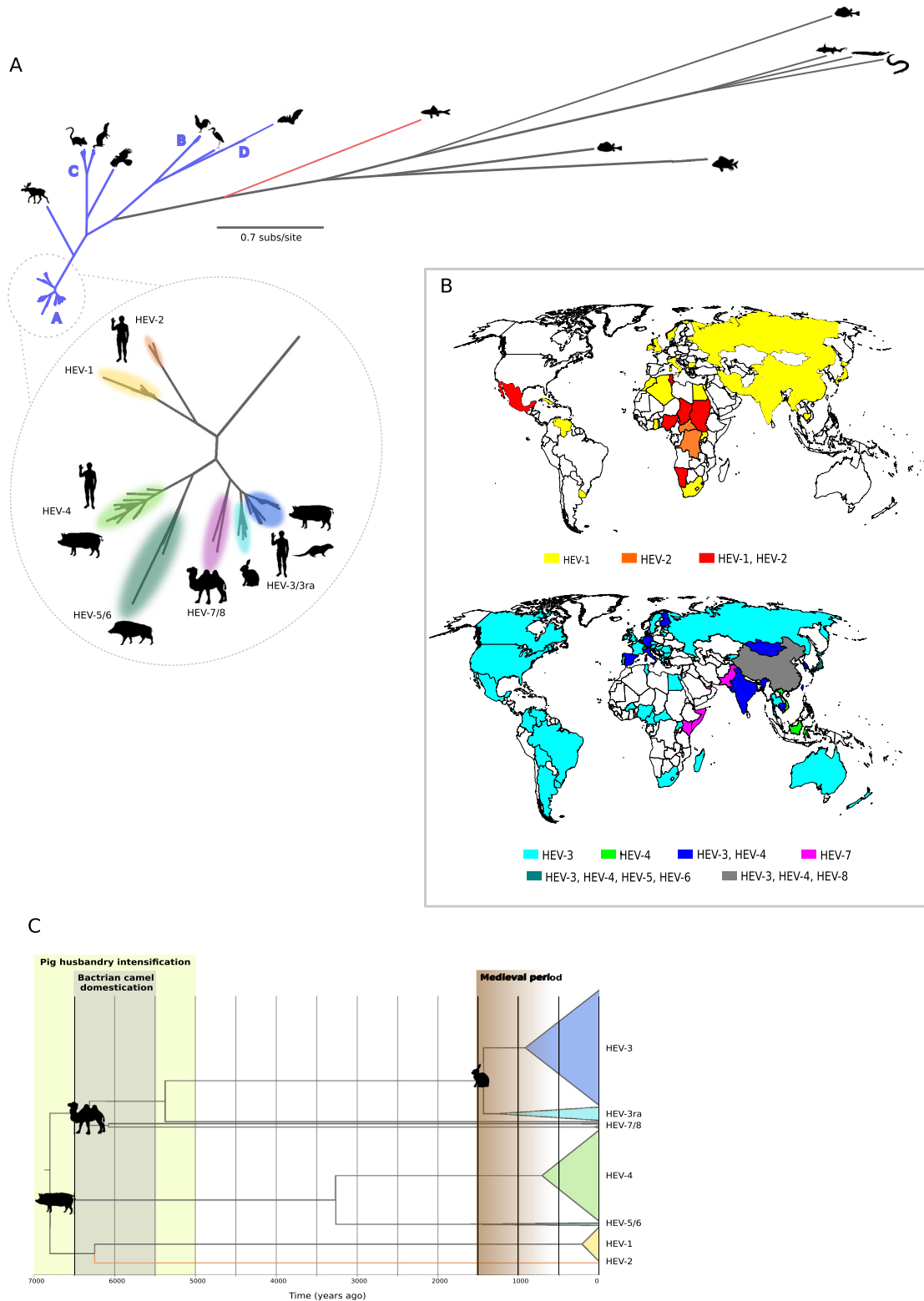
HDV is a defective virus incapable of autonomous propagation [93]. Its genome, a self complementary circular RNA of ~1700 nucleotides, encodes a single protein (the HDV-encoded delta antigen) (Table 1). HDV requires HBV surface proteins, that are complexed with the delta antigen, to form transmissible virions [94]. Thus, HDV is usually considered a satellite of HBV, although recent data have shown that other enveloped viruses can promote HDV propagation, at least *in vitro* (e.g. HCV, dengue virus, vesicular stomatitis virus) [95].

Genetic heterogeneity among strains is quite high for HDV, which is thus classified in eight different genotypes (from 1 to 8), although a three major genogroup classification was recently proposed (group 1 for previous genotype 1, group 2 for genotypes 2 and genotypes from 4 to 8, and group 3 for previous genotype 3) [96]. HDV genetic diversity is highest in Africa, suggesting that the defective virus emerged in and spread from this continent [40,97]. The evolutionary origin of HDV is nonetheless unknown. HDV-like circular RNAs were only recently described in birds, reptiles, amphibians, fish and insects [98–100]. However, these HDV-like elements were not found to be associated with hepadnavirus infection, reinforcing the idea the HDV is not necessarily only transmitted in conjunction with HBV-related viruses, at least in these animals [100]. Current evidence suggests that HDV evolved from the human cellular transcriptome [101].

## 7. Hepatitis E virus

The first indications that a virus was responsible for waterborne, epidemic hepatitis came from studies of Asian outbreaks in the 1950–70s and, in analogy to HCV, the agent was referred to as “epidemic non-A, non-B hepatitis” [102–104]. Hepatitis E virus was eventually isolated and sequenced in the early 1990s [105,106]. Since then, a number of HEV strains responsible for human infection were identified.

HEV is a positive-strand RNA virus belonging to the *Hepeviridae* family (Table 1). In common with all other members of this viral family, the HEV genome comprises three partially overlapping open reading frames (ORFs): ORF1 and ORF2 encode a non-structural



**Fig. 4.** Hepevirus diversity and geographic distribution. (A) A maximum-likelihood phylogeny of the RdRp domain of known hepeviruses was generated with PhyML [136]. The Piscihepevirus branch is in red. Orthohepevirus branches are in blue. The enlargement shows phylogenetic relationships for viruses belonging to the *Orthohepevirus A* species, with representative hosts. (B) Geographic distribution of anthrotopropic (HEV-1 and HEV-2) and enzootic (HEV-3–HEV-4) HEV strains. Genotypes were assigned to countries irrespective of their prevalence. Thus, even if a single case was reported in a given country, the genotype was recorded as present. Cases that could be clearly ascribed to migration/travels were excluded. Data derive from Forni et al. [121], with updates from [110,140–151]. (C) Time-scaled phylogenetic tree of a non-recombining ORF1 region [121]. Branch lengths represent evolutionary time. The time-frames of historical events mentioned in the text are reported. The rabbit and camel silhouettes mark the split of the rabbit-infecting and camel/dromedary-infecting genotypes. The pig silhouette marks the human-restricted/enzootic genotype split.



polyprotein and the viral capsid, respectively, whereas ORF3 codes for a small phosphoprotein. A fourth ORF (ORF4), overlapping the helicase domain in ORF1, was recently described but seems to be specific for some HEV genotypes (HEV genotype 1, HEV-1) [107].

Members of the *Hepeviridae* family are currently classified into two genera, *Orthohepevirus* and *Piscihepevirus* [108] ([https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/positive-sense-rna-viruses/w/hepeviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/hepeviridae)). The *Piscihepevirus* genus includes only one species (*Piscihepevirus A*) with one member (cutthroat trout virus), whereas the *Orthohepevirus* genus is divided into four species of viruses infecting mammals and birds (*Orthohepevirus A–D*) [108] (Fig. 4A). This classification is, however, likely to change in the near future following the identification of novel hepeviruses in fish other than trout and in amphibians [28] (Fig. 4A).

Human-infecting HEV strains are genetically heterogeneous and display distinct epidemiologic patterns, but all belong to the *Orthohepevirus A* species. *Orthohepeviruses A* account for a minority of the overall diversity of hepeviruses that infect vertebrates and their closest relative is a presently unclassified virus detected in a Swedish moose (Fig. 4A) [109]. Field surveys revealed a high prevalence of HEV in moose populations from Sweden and other Baltic Regions [110,111]. In general, ungulates represent major *Orthohepeviruses A* reservoirs. At present, eight *Orthohepevirus A* genotypes are recognized (HEV-1 to -8) (Fig. 4A). HEV-1 and HEV-2 infect only humans and cause waterborne outbreaks mainly in tropical and subtropical regions (Fig. 4B). Conversely, genotypes 3 and 4 account for the majority of hepatitis E human cases in industrialized countries. HEV-3 and HEV-4 also infect several other domestic (mainly pigs) and wild (e.g., ungulates and small carnivores) animals, their transmission to humans being usually zoonotic [2] (Fig. 4A). Phylogenetic analyses showed that HEV-3 and HEV-4 sequences derived from human cases are interspersed within those isolated from swine, indicating that pig-infecting HEV-3 and HEV-4 can easily cross the species barrier and infect humans [112]. Notably, though, evolutionary rates are higher for genotypes 3 and 4 than for the human-specific genotypes, suggesting cyclical adaptation to different mammalian hosts [113]. A distinct HEV-3 clade, mainly detected in rabbits (HEV-3ra), can also cause human hepatitis E [114–117] (Fig. 4A). The remaining genotypes HEV-5/HEV-6 and HEV-7/HEV-8 have been detected in boars and camels, respectively [2] (Fig. 4A). However, they are also thought to have zoonotic potential, as HEV-5 and HEV-7 can infect cynomolgus monkeys [118,119] and HEV-7 was detected in a patient who consumed camel meat and milk [120]. Thus, viruses belonging to all HEV genotypes seem to be transmissible to humans. Conversely, experimental infection with HEV-1 and HEV-2 indicated that these viruses have a host range restricted to primates [2]. HEV genotypes are therefore usually referred to as enzootic (HEV-3 and -4) or human-restricted/anthropotropic (HEV-1 and -2).

Although several human hepatitis E cases have a zoonotic origin and *orthohepeviruses A* are found in diverse mammals, recent data indicated that one or more reverse zoonoses led to the emergence and radiation of HEV genotypes [121]. In fact, character state reconstruction on a large phylogeny revealed that humans were the most likely hosts of the ancestor of extant *orthohepeviruses A* [121]. This notion is in line with the observation that most, if not all, HEV genotypes can infect our species, whereas other animals are differentially susceptible to distinct HEV genotypes. Moreover, increasing evidence suggests that reverse zoonotic events (also known as zoonothronoses) are all but rare, and examples include other RNA viruses such as rotaviruses, enteroviruses, and human influenza viruses [122–125]. For both swine influenza A viruses and swine vesicular disease virus onward transmission in pigs is well documented [123,125] and is facilitated by intensive husbandry practices.

Molecular dating using a selection-informed method inferred that the ancestor of extant *orthohepeviruses A* existed ~6800 to ~3200 years ago, most likely in East Asia [121]. These inferences well correlate with historical circumstances that may have favored HEV emergence and host range expansion. In this period, sedentary agriculture promoted the appearance of large human settlements in several Asian regions and pig husbandry practices started to intensify in East Asia [126–130] (Fig. 4C). Crowded living conditions and poor sanitation possibly favored the emergence and spread of the waterborne, human-specific HEV strains. The close contact between humans and pigs most likely promoted HEV zoonothronotic transmission and emergence of the enzootic strains (Fig. 4C). Additional reverse zoonotic transmissions may have also originated the camel-infecting and rabbit-infecting strains. In fact, the estimated timing of HEV-7/8 emergence (4055 BCE to 1192 BCE) [121] encompasses the time of domestication of Bactrian and dromedary camels [131–133] (Fig. 4C). As for HEV-3ra, it was estimated to have diverged from HEV-3 around 600 CE, in Europe [121]. This time frame corresponds to the Middle Ages, when historical evidence indicates that rabbits were kept in warrens and bred for meat [134] (Fig. 4C). Of course, these estimates do not necessarily imply that camels and rabbits acquired HEV from humans, as the domestication process may have exposed these animals to various viral sources, including other domestic (e.g., pigs) and peridomestic mammals.

These scenarios provide a credible framework for *orthohepevirus A* origin, as well as for the diversification of HEV genotypes, and selection-informed methods should at least partially correct for the TDRP, as they explicitly model purifying selection [14–16]. Nonetheless, the above-mentioned data on HBV [9] suggest caution in the inference of time and location of ancestral events based on extant viral diversity. Also, pig infection with HEV-3 and HEV-4 is generally asymptomatic [135], possibly indicating a long-standing host-virus association that might even predate swine husbandry development.

It should also be noted that the ultimate origin of *orthohepevirus A* remains unknown. Humans may have acquired HEV through cross-species transmission from other animals. However, known *orthohepeviruses* that infect mammals and birds are distantly related to *orthohepevirus A*, indicating that none of them represents the source of human-infecting HEV. Likewise, the origin and evolutionary relationship between the moose virus and human-infecting *orthohepeviruses* remain unclear.

## 8. Conclusions

By allowing the large-scale identification of novel viral species, as well as the sequencing of viral genomes from archaeological samples, technological advances have largely expanded our knowledge on the evolution and origin of human hepatitis viruses. These insights have been paralleled by the development of computational tools and theoretical frameworks to analyze and mine viral sequence data (e.g., the above-mentioned recognition of the TDRP and the development of approaches to correct for it). The overall picture emerging from these studies clearly indicates that, with the possible exception of HDV, viruses related to human hepatitis viruses infect several other mammalian and non mammalian vertebrates. The specific events that originated these human pathogens remain, however, largely unknown. For HCV and HEV, the evolutionary history of the human viruses is likely intertwined with that of related viruses that infect domestic animals. Conversely, in the case of HAV and HBV, the closest relatives of the human viruses are found in NHPs. In general, these observations indicate that a deeper understanding of the evolutionary dynamics of human viruses has a relevance not only to improve our ability to

treat and prevent present infections, but also to gain insight into the ecological contexts that may fuel the emergence of novel human pathogens, with particular reference to zoonotic ones.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgement

This work was supported by the Italian Ministry of Health, grant n. RC 2019 to MS, grant n. RC 2018–2019 to DF.

### References

- World Health Organization (WHO). Global hepatitis report 2017. <https://www.who.int/hepatitis/publications/global-hepatitis-report2017/en/>.
- Doceul V, Bagdassarian E, Demange A, Pavio N. Zoonotic hepatitis E virus: classification, animal reservoirs and transmission routes. *Viruses* 2016;8:E270.
- Cuthbert JA. Hepatitis A: old and new. *Clin Microbiol Rev* 2001;14:38–58.
- Lemon SM, Ott JJ, Van Damme P, Shouval D. Type A viral hepatitis: a summary and update on the molecular virology, epidemiology, pathogenesis and prevention. *J Hepatol* 2017.
- Bromham L, Duchene S, Hua X, Ritchie AM, Duchene DA, Ho SYW. Bayesian molecular dating: opening up the black box. *Biol Rev Camb Philos Soc* 2018;93:1165–91.
- Rieux A, Balloux F. Inferences from tip-calibrated phylogenies: a review and a practical guide. *Mol Ecol* 2016;25:1911–24.
- Zhou Y, Holmes EC. Bayesian estimates of the evolutionary rate and age of hepatitis B virus. *J Mol Evol* 2007;65:197–205.
- Zehender G, Ebranati E, Gabanelli E, Sorrentino C, Lo Presti A, Tanzi E, et al. Enigmatic origin of hepatitis B virus: an ancient travelling companion or a recent encounter? *World J Gastroenterol* 2014;20:7622–34.
- Muhlemann B, Jones TC, Damgaard PB, Allentoft ME, Shevina I, Logvin A, et al. Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature* 2018;557:418–23.
- Aiewsakun P, Katzourakis A. Time-dependent rate phenomenon in viruses. *J Virol* 2016;90:7184–95.
- Simmonds P, Aiewsakun P, Katzourakis A. Prisoners of war – host adaptation and its constraints on virus evolution. *Nat Rev Microbiol* 2019;17:321–8.
- Ho SY, Duchene S, Molak M, Shapiro B. Time-dependent estimates of molecular evolutionary rates: evidence and causes. *Mol Ecol* 2015;24:6007–12.
- Membrebe JV, Suchard MA, Rambaut A, Baele G, Lemey P. Bayesian inference of evolutionary histories under time-dependent substitution rates. *Mol Biol Evol* 2019.
- Wertheim JO, Chu DK, Peiris JS, Kosakovsky Pond SL, Poon LL. A case for the ancient origin of coronaviruses. *J Virol* 2013;87:7039–45.
- Wertheim JO, Kosakovsky Pond SL. Purifying selection can obscure the ancient age of viral lineages. *Mol Biol Evol* 2011;28:3355–65.
- Wertheim JO, Smith MD, Smith DM, Scheffler K, Kosakovsky Pond SL. Evolutionary origins of human herpes simplex viruses 1 and 2. *Mol Biol Evol* 2014;31:2356–64.
- McKnight KL, Lemon SM. Hepatitis A virus genome organization and replication strategy. *Cold Spring Harb Perspect Med* 2018;8. <https://doi.org/10.1101/cshperspect.a033480>.
- Feng Z, Hensley L, McKnight KL, Hu F, Madden V, Ping L, et al. A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. *Nature* 2013;496:367–71.
- Aragones L, Guix S, Ribes E, Bosch A, Pinto RM. Fine-tuning translation kinetics selection as the driving force of codon usage bias in the hepatitis A virus capsid. *PLoS Pathog* 2010;6:e1000797.
- Feinstone SM, Kapikian AZ, Purcell RH. Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. *Science* 1973;182:1026–8.
- Skinhoj P, Mikkelsen F, Hollinger FB. Hepatitis A in Greenland: importance of specific antibody testing in epidemiologic surveillance. *Am J Epidemiol* 1977;105:140–7.
- Wong DC, Purcell RH, Rosen L. Prevalence of antibody to hepatitis A and hepatitis B viruses in selected populations of the South Pacific. *Am J Epidemiol* 1979;110:227–36.
- Cristina J, Costa-Mattioli M. Genetic variability and molecular evolution of hepatitis A virus. *Virus Res* 2007;127:151–7.
- Drexler JF, Corman VM, Lukashev AN, van den Brand JM, Gmyl AP, Brunink S, et al. Evolutionary origins of hepatitis A virus in small mammals. *Proc Natl Acad Sci USA* 2015;112:15190–5.
- Anthony SJ, St Leger JA, Liang E, Hicks AL, Sanchez-Leon MD, Jain K, et al. Discovery of a novel hepatovirus (phopivirus of seals) related to human hepatitis A virus. *MBio* 2015;6. <https://doi.org/10.1128/mBio.01180-15>.
- Yu JM, Li LL, Zhang CY, Lu S, Ao YY, Gao HC, et al. A novel hepatovirus identified in wild woodchuck *Marmota himalayana*. *Sci Rep* 2016;6:22361.
- de Oliveira Carneiro I, Sander AL, Silva N, Moreira-Soto A, Normann A, Flehming B, et al. A novel marsupial hepatitis A virus corroborates complex evolutionary patterns shaping the genus hepatovirus. *J Virol* 2018;92. <https://doi.org/10.1128/JVI.00082-18>. Print 2018 Jul 1.
- Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, et al. The evolutionary history of vertebrate RNA viruses. *Nature* 2018;556:197–202.
- Sander AL, Corman VM, Lukashev AN, Drexler JF. Evolutionary origins of enteric hepatitis viruses. *Cold Spring Harb Perspect Med* 2018;8. <https://doi.org/10.1101/cshperspect.a031690>.
- Guo WP, Lin XD, Wang W, Tian JH, Cong ML, Zhang HL, et al. Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. *PLoS Pathog* 2013;9:e1003159.
- Wang X, Ren J, Gao Q, Hu Z, Sun Y, Li X, et al. Hepatitis A virus and the origins of picornaviruses. *Nature* 2015;517:85–8.
- Dane DS, Cameron CH, Briggs M. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet* 1970;1:695–8.
- MacLachlan JH, Cowie BC. Hepatitis B virus epidemiology. *Cold Spring Harb Perspect Med* 2015;5:a021410.
- Caballero A, Taberno D, Buti M, Rodriguez-Frias F. Hepatitis B virus: the challenge of an ancient virus with multiple faces and a remarkable replication strategy. *Antiviral Res* 2018;158:34–44.
- Kramvis A. Genotypes and genetic variability of hepatitis B virus. *Intervirology* 2014;57:141–50.
- Castelhano N, Araujo NM, Arenas M. Heterogeneous recombination among Hepatitis B virus genotypes. *Infect Genet Evol* 2017;54:486–90.
- Tatematsu K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, et al. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. *J Virol* 2009;83:10538–47.
- Locarnini S, Littlejohn M, Aziz MN, Yuen L. Possible origins and evolution of the hepatitis B virus (HBV). *Semin Cancer Biol* 2013;23:561–75.
- Velkov S, Ott JJ, Protzer U, Michler T. The global hepatitis B virus genotype distribution approximated from available genotyping data. *Genes (Basel)* 2018;9. <https://doi.org/10.3390/genes9100495>.
- Littlejohn M, Locarnini S, Yuen L. Origins and evolution of hepatitis B virus and hepatitis D virus. *Cold Spring Harb Perspect Med* 2016;6:a021360.
- de Carvalho Dominguez Souza BF, König A, Rasche A, de Oliveira Carneiro I, Stephan N, Corman VM, et al. A novel hepatitis B virus species discovered in capuchin monkeys sheds new light on the evolution of primate hepadnaviruses. *J Hepatol* 2018;68:1114–22.
- Rasche A, Souza BFC, Drexler JF. Bat hepadnaviruses and the origins of primate hepatitis B viruses. *Curr Opin Virol* 2016;16:86–94.
- Lauber C, Seitz S, Mattei S, Suh A, Beck J, Herstein J, et al. Deciphering the origin and evolution of hepatitis B viruses by means of a family of non-enveloped fish viruses. *Cell Host Microbe* 2017;22(387–399):e6.
- Gilbert C, Meik JM, Dashevsky D, Card DC, Castoe TA, Schaack S. Endogenous hepadnaviruses, bornaviruses and circoviruses in snakes. *Proc Biol Sci* 2014;281:20141122.
- Suh A, Weber CC, Kehlmaier C, Braun EL, Green RE, Fritz U, et al. Early mesozoic coexistence of amniotes and hepadnaviridae. *PLoS Genet* 2014;10:e1004559.
- Cui J, Holmes EC. Endogenous hepadnaviruses in the genome of the budgerigar (*Melopsittacus undulatus*) and the evolution of avian hepadnaviruses. *J Virol* 2012;86:7688–91.
- Liu W, Pan S, Yang H, Bai W, Shen Z, Liu J, et al. The first full-length endogenous hepadnaviruses: identification and analysis. *J Virol* 2012;86:9510–3.
- Bollyky P, Rambaut A, Grassley N, Carman W, Holmes E. Hepatitis B virus has a recent new world evolutionary origin. *J Hepatol* 1998;28:96.
- Simmonds P. Reconstructing the origins of human hepatitis viruses. *Philos Trans R Soc Lond B Biol Sci* 2001;356:1013–26.
- Kahila Bar-Gal G, Kim MJ, Klein A, Shin DH, Oh CS, Kim JW, et al. Tracing hepatitis B virus to the 16th century in a Korean mummy. *Hepatology* 2012;56:1671–80.
- Norder H, Courouze AM, Magnius LO. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* 1994;198:489–503.
- Magnius LO, Norder H. Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. *Intervirology* 1995;38:24–34.
- Paraskevis D, Magiorkinis G, Magiorkinis E, Ho SY, Belshaw R, Allain JP, et al. Dating the origin and dispersal of hepatitis B virus infection in humans and primates. *Hepatology* 2013;57:908–16.
- Paraskevis D, Angelis K, Magiorkinis G, Kostaki E, Ho SY, Hatzakis A. Dating the origin of hepatitis B virus reveals higher substitution rate and adaptation on the branch leading to F/H genotypes. *Mol Phylogenet Evol* 2015;93:44–54.
- Krause-Kyora B, Susat J, Key FM, Kuhnert D, Bosse E, Immel A, et al. Neolithic and medieval virus genomes reveal complex evolution of hepatitis. *BELIFE* 2018;7. <https://doi.org/10.7554/eLife.36666>.
- MacDonald DM, Holmes EC, Lewis JC, Simmonds P. Detection of hepatitis B virus infection in wild-born chimpanzees (*Pan troglodytes verus*):

- phylogenetic relationships with human and other primate genotypes. *J Virol* 2000;74:4253–7.
- [57] Patterson Ross Z, Klunk J, Fornaciari G, Giuffra V, Duchene S, Duggan AT, et al. The paradox of HBV evolution as revealed from a 16th century mummy. *PLoS Pathog* 2018;14:e1006750.
- [58] Forbi JC, Vaughan G, Purdy MA, Campo DS, Xia GL, Ganova-Raeva LM, et al. Epidemic history and evolutionary dynamics of hepatitis B virus infection in two remote communities in rural Nigeria. *PLoS One* 2010;5:e11615.
- [59] Mojsiejczuk LN, Torres C, Sevic I, Badano I, Malan R, Flichman DM, et al. Molecular epidemiology of hepatitis B virus in Misiones, Argentina. *Infect Genet Evol* 2016;44:34–42.
- [60] Sagnelli C, Ciccozzi M, Coppola N, Minichini C, Lo Presti A, Starace M, et al. Molecular diversity in irregular or refugee immigrant patients with HBV-genotype-E infection living in the metropolitan area of Naples. *J Med Virol* 2017;89:1015–24.
- [61] Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989;244:359–62.
- [62] Jackowiak P, Kuls K, Budzko L, Mania A, Figlerowicz M, Figlerowicz M. Phylogeny and molecular evolution of the hepatitis C virus. *Infect Genet Evol* 2014;21:67–82.
- [63] Pybus OG, Charleston MA, Gupta S, Rambaut A, Holmes EC, Harvey PH. The epidemic behavior of the hepatitis C virus. *Science* 2001;292:2323–5.
- [64] Simmonds P. The origin of hepatitis C virus. *Curr Top Microbiol Immunol* 2013;369:1–15.
- [65] Ray SC, Arthur RR, Carella A, Bukh J, Thomas DL. Genetic epidemiology of hepatitis C virus throughout Egypt. *J Infect Dis* 2000;182:698–707.
- [66] Messina JP, Humphreys I, Flaxman A, Brown A, Cooke GS, Pybus OG, et al. Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology* 2015;61:77–87.
- [67] Borgia SM, Hedskog C, Parhy B, Hyland RH, Stamm LM, Brainard DM, et al. Identification of a novel hepatitis C virus genotype from Punjab, India: expanding classification of hepatitis C virus into 8 genotypes. *J Infect Dis* 2018;218:1722–9.
- [68] Hedskog C, Parhy B, Chang S, Zeuzem S, Moreno C, Shafraan SD, et al. Identification of 19 novel hepatitis C virus subtypes-further expanding HCV classification. *Open Forum Infect Dis* 2019;6:ofz076.
- [69] Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis* 2005;5:558–67.
- [70] Forni D, Cagliani R, Pontremoli C, Pozzoli U, Vertemara J, De Gioia L, et al. Evolutionary analysis provides insight into the origin and adaptation of HCV. *Front Microbiol* 2018;9:854.
- [71] Chan DP, Sun HY, Wong HT, Lee SS, Hung CC. Sexually acquired hepatitis C virus infection: a review. *Int J Infect Dis* 2016;49:47–58.
- [72] Dobson AP, Carper ER. Infectious diseases and human population history. *Bioscience* 1996;46:115–26.
- [73] Pybus OG, Markov PV, Wu A, Tatem AJ. Investigating the endemic transmission of the hepatitis C virus. *Int J Parasitol* 2007;37:839–49.
- [74] Houldsworth A. Exploring the possibility of arthropod transmission of HCV. *J Med Virol* 2017;89:187–94.
- [75] Corman VM, Grundhoff A, Baechlein C, Fischer N, Gmyl A, Wollny R, et al. Highly divergent hepaciviruses from African cattle. *J Virol* 2015;89:5876–82.
- [76] Kapoor A, Simmonds P, Gerold G, Qaisar N, Jain K, Henriquez JA, et al. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci USA* 2011;108:11608–13.
- [77] Kapoor A, Simmonds P, Scheel TK, Hjelle B, Cullen JM, Burbelo PD, et al. Identification of rodent homologs of hepatitis C virus and pegiviruses. *MBio* 2013;4:e00216–13.
- [78] Scheel TK, Simmonds P, Kapoor A. Surveying the global virome: identification and characterization of HCV-related animal hepaciviruses. *Antiviral Res* 2015;115:83–93.
- [79] Drexler JF, Corman VM, Muller MA, Lukashev AN, Gmyl A, Coutard B, et al. Evidence for novel hepaciviruses in rodents. *PLoS Pathog* 2013;9:e1003438.
- [80] Quan PL, Firth C, Conte JM, Williams SH, Zambrana-Torrel CM, Anthony SJ, et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses. *Proc Natl Acad Sci USA* 2013;110:8194–9.
- [81] Burbelo PD, Dubovi EJ, Simmonds P, Medina JL, Henriquez JA, Mishra N, et al. Serology-enabled discovery of genetically diverse hepaciviruses in a new host. *J Virol* 2012;86:6171–8.
- [82] Walter S, Rasche A, Moreira-Soto A, Pfaender S, Bletsch M, Corman VM, et al. Differential infection patterns and recent evolutionary origins of equine hepaciviruses in donkeys. *J Virol* 2016;91. Print 2017 Jan 1e01711–16.
- [83] Baechlein C, Fischer N, Grundhoff A, Alawi M, Indenbirken D, Postel A, et al. Identification of a novel hepacivirus in domestic cattle from Germany. *J Virol* 2015;89:7007–15.
- [84] El-Attar LM, Mitchell JA, Brooks Brownlie H, Priestnall SL, Brownlie J. Detection of non-primate hepaciviruses in UK dogs. *Virology* 2015;484:93–102.
- [85] Pybus OG, Theze J. Hepacivirus cross-species transmission and the origins of the hepatitis C virus. *Curr Opin Virol* 2016;16:1–7.
- [86] Gray RR, Tanaka Y, Takebe Y, Magiorikinis G, Buskell Z, Seeff L, et al. Evolutionary analysis of hepatitis C virus gene sequences from 1953. *Philos Trans R Soc Lond B Biol Sci* 2013;368:20130168.
- [87] Li C, Lu L, Murphy DG, Negro F, Okamoto H. Origin of hepatitis C virus genotype 3 in Africa as estimated through an evolutionary analysis of the full-length genomes of nine subtypes, including the newly sequenced 3d and 3e. *J Gen Virol* 2014;95:1677–88.
- [88] Lu L, Li C, Xu Y, Murphy DG. Full-length genomes of 16 hepatitis C virus genotype 1 isolates representing subtypes 1c, 1d, 1e, 1g, 1h, 1i, 1j and 1k, and two new subtypes 1m and 1n, and four unclassified variants reveal ancestral relationships among subtypes. *J Gen Virol* 2014;95:1479–87.
- [89] Smith DB, Pathirana S, Davidson F, Lawlor E, Power J, Yap PL, et al. The origin of hepatitis C virus genotypes. *J Gen Virol* 1997;78(Pt 2):321–8.
- [90] Muhlemann B, Margaryan A, Damgaard PB, Allentoft ME, Vinner L, Hansen AJ, et al. Ancient human parvovirus B19 in Eurasia reveals its long-term association with humans. *Proc Natl Acad Sci USA* 2018;115:7557–62.
- [91] Duggan AT, Perdomo MF, Piombino-Mascalci D, Marciniak S, Poinar D, Emery MV, et al. 17(th) Century variola virus reveals the recent history of smallpox. *Curr Biol* 2016;26:3407–12.
- [92] Fages A, Hanghoj K, Khan N, Gaunitz C, Seguin-Orlando A, Leonardi M, et al. Tracking five millennia of horse management with extensive ancient genome time series. *Cell* 2019;177(1419–1435):e31.
- [93] Rizzetto M, Canese MG, Gerin JL, London WT, Sly DL, Purcell RH. Transmission of the hepatitis B virus-associated delta antigen to chimpanzees. *J Infect Dis* 1980;141:590–602.
- [94] Hughes SA, Wedemeyer H, Harrison PM. Hepatitis delta virus. *Lancet* 2011;378:73–85.
- [95] Perez-Vargas J, Amirache F, Boson B, Mialon C, Freitas N, Sureau C, et al. Enveloped viruses distinct from HBV induce dissemination of hepatitis D virus in vivo. *Nat Commun* 2019;10. 2098–019-10117-z.
- [96] Delfino CM, Cerrudo CS, Biglione M, Oubina JR, Ghiringhelli PD, Mathet VL. A comprehensive bioinformatic analysis of hepatitis D virus full-length genomes. *J Viral Hepat* 2018;25:860–9.
- [97] Le Gal F, Brichler S, Drugan T, Alloui C, Roulot D, Pawlotsky JM, et al. Genetic diversity and worldwide distribution of the deltavirus genus: a study of 2,152 clinical strains. *Hepatology* 2017;66:1826–41.
- [98] Wille M, Netter HJ, Littlejohn M, Yuen L, Shi M, Eden JS, et al. A divergent hepatitis D-like agent in birds. *Viruses* 2018;10. 10.3390/v10120720.
- [99] Hetzel U, Szirovicza L, Smura T, Prahauer B, Vapalahti O, Kipar A, et al. Identification of a novel deltavirus in boa constrictors. *MBio* 2019;10. <https://doi.org/10.1128/mBio.00014-19>.
- [100] Chang W, Pettersson JH, Le Lay C, Shi M, Lo N, Wille M, et al. Novel hepatitis D-like agents in vertebrates and invertebrates. *ve* 2019;5.
- [101] Cunha C, Tavanez JP, Gudima S. Hepatitis delta virus: a fascinating and neglected pathogen. *World J Virol* 2015;4:313–22.
- [102] Khuroo MS. Study of an epidemic of non-A, non-B hepatitis. possibility of another human hepatitis virus distinct from post-transfusion non-A, non-B type. *Am J Med* 1980;68:818–24.
- [103] Viswanathan R. Infectious hepatitis in Delhi (1955–56): a critical study-epidemiology. 1957. *Natl Med J India* 2013;26:362–77.
- [104] Wong DC, Purcell RH, Sreenivasan MA, Prasad SR, Pavri KM. Epidemic and endemic hepatitis in India: evidence for a non-A, non-B hepatitis virus aetiology. *Lancet* 1980;2:876–9.
- [105] Reyes GR, Purdy MA, Kim JP, Luk KC, Young LM, Fry KE, et al. Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis. *Science* 1990;247:1335–9.
- [106] Tam AW, Smith MM, Guerra ME, Huang CC, Bradley DW, Fry KE, et al. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. *Virology* 1991;185:120–31.
- [107] Nair VP, Anang S, Subramani C, Madhvi A, Bakshi K, Srivastava A, et al. Endoplasmic reticulum stress induced synthesis of a novel viral factor mediates efficient replication of genotype-1 hepatitis E virus. *PLoS Pathog* 2016;12:e1005521.
- [108] Smith DB, Simmonds P, International Committee on Taxonomy of Viruses Hepeviridae Study Group, Jameel S, Emerson SU, Harrison TJ, et al. Consensus proposals for classification of the family Hepeviridae. *J Gen Virol* 2014;95:2223–32.
- [109] Lin J, Norder H, Uhlhorn H, Belak S, Widen F. Novel hepatitis E like virus found in Swedish moose. *J Gen Virol* 2014;95:557–70.
- [110] Spanceniene U, Grigas J, Buitkuviene J, Zymantiene J, Juozaitiene V, Stankeviciute M, et al. Prevalence and phylogenetic analysis of hepatitis E virus in pigs, wild boars, roe deer, red deer and moose in Lithuania. *Acta Vet Scand* 2018;60. 13–018-0367-7.
- [111] Lin J, Karlsson M, Olofson AS, Belak S, Malmsten J, Dalin AM, et al. High prevalence of hepatitis E virus in Swedish moose – a phylogenetic characterization and comparison of the virus from different regions. *PLoS One* 2015;10:e0122102.
- [112] Bouquet J, Tesse S, Lunazzi A, Eloit M, Rose N, Nicand E, et al. Close similarity between sequences of hepatitis E virus recovered from humans and swine, France, 2008–2009. *Emerg Infect Dis* 2011;17:2018–25.
- [113] Brayne AB, Dearlove BL, Lester JS, Kosakovsky Pond SL, Frost SD. Genotype-specific evolution of hepatitis E virus. *J Virol* 2017;91. <https://doi.org/10.1128/JVI.02241-16>. Print 2017 May 1.
- [114] Abravanel F, Lhomme S, El Costa H, Schvartz B, Peron JM, Kamar N, et al. Rabbit hepatitis E virus infections in humans, France. *Emerg Infect Dis* 2017;23:1191–3.
- [115] Sahli R, Fraga M, Semela D, Moradpour D, Gouttenoire J. Rabbit HEV in immunosuppressed patients with hepatitis E acquired in Switzerland. *J Hepatol* 2019;70:1023–5.
- [116] Liu P, Bu QN, Wang L, Han J, Du RJ, Lei YX, et al. Transmission of hepatitis E virus from rabbits to cynomolgus macaques. *Emerg Infect Dis* 2013;19:559–65.

- [117] Izopet J, Dubois M, Bertagnoli S, Lhomme S, Marchandeau S, Boucher S, et al. Hepatitis E virus strains in rabbits and evidence of a closely related strain in humans, France. *Emerg Infect Dis* 2012;18:1274–81.
- [118] Li TC, Bai H, Yoshizaki S, Ami Y, Suzaki Y, Doan YH, et al. Genotype 5 hepatitis E virus produced by a reverse genetics system has the potential for zoonotic infection. *Hepatal Commun* 2018;3:160–72.
- [119] Li TC, Zhou X, Yoshizaki S, Ami Y, Suzaki Y, Nakamura T, et al. Production of infectious dromedary camel hepatitis E virus by a reverse genetic system: potential for zoonotic infection. *J Hepatol* 2016;65:1104–11.
- [120] Lee GH, Tan BH, Teo EC, Lim SG, Dan YY, Wee A, et al. Chronic infection with camelid hepatitis E virus in a liver transplant recipient who regularly consumes camel meat and milk. *Gastroenterology* 2016;150(355–7):e3.
- [121] Forni D, Cagliani R, Clerici M, Sironi M. Origin and dispersal of hepatitis E virus. *Emerg Microbes Infect* 2018;7: 11-017-0009-6.
- [122] Messenger AM, Barnes AN, Gray GC. Reverse zoonotic disease transmission (zooanthroponosis): a systematic review of seldom-documented human biological threats to animals. *PLoS One* 2014;9:e89055.
- [123] Nelson MI, Vincent AL. Reverse zoonosis of influenza to swine: new perspectives on the human-animal interface. *Trends Microbiol* 2015;23:142–53.
- [124] Fieldhouse JK, Wang X, Mallinson KA, Tsao RW, Gray GC. A systematic review of evidence that enteroviruses may be zoonotic. *Emerg Microbes Infect* 2018;7: 164-018-0159-1.
- [125] Huang HW, Chu PH, Pan CH, Wang CF, Lin CC, Lu PL, et al. Evolutionary histories of coxsackievirus B5 and swine vesicular disease virus reconstructed by phylodynamic and sequence variation analyses. *Sci Rep* 2018;8: 8821-018-27254-y.
- [126] Bocquet-Appel J. Bar-Yosef of the Neolithic demographic transition and its consequences, 2008.
- [127] Gignoux CR, Henn BM, Mountain JL. Rapid, global demographic expansions after the origins of agriculture. *Proc Natl Acad Sci USA* 2011;108:6044–9.
- [128] Underhill AP, Feinman GM, Nicholas LM, Fang H, Luan F, Yu H, et al. Changes in regional settlement patterns and the development of complex societies in southeastern Shandong, China. *J Anthropol Archaeol* 2008;27:1–29.
- [129] Cucchi T, Dai L, Balasse M, Zhao C, Gao J, Hu Y, et al. Social complexification and pig (*Sus scrofa*) husbandry in ancient China: a combined geometric morphometric and isotopic approach. *PLoS One* 2016;11:e0158523.
- [130] Barton L, Newsome SD, Chen FH, Wang H, Guilderson TP, Bettinger RL. Agricultural origins and the isotopic identity of domestication in northern China. *Proc Natl Acad Sci USA* 2009;106:5523–8.
- [131] Kuzmina EE. The prehistory of the silk road, 2008.
- [132] Grigson C. The history of the camel bone dating project. *Anthropozoologica* 2014;49:225–35.
- [133] Uerpmann H, Uerpmann M. The appearance of the domestic camel in south-east Arabia. *J Oman Stud* 2002;12:235–60.
- [134] Irving-Pease EK, Frantz LAF, Sykes N, Callou C, Larson G. Rabbits and the specious origins of domestication. *Trends Ecol Evol* 2018;33:149–52.
- [135] Pavio N, Doceul V, Bagdassarian E, Johne R. Recent knowledge on hepatitis E virus in Suidae reservoirs and transmission routes to human. *Vet Res* 2017;48: 78-017-0483-9.
- [136] Guindon S, Delsuc F, Dufayard JF, Gascuel O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* 2009;537:113–37.
- [137] Alvarado-Mora MV, Pinho JR. Distribution of HBV genotypes in Latin America. *Antivir Ther* 2013;18:459–65.
- [138] Polaris Observatory HCV Collaborators. Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study. *Lancet Gastroenterol Hepatol* 2017;2:161–76.
- [139] Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–3.
- [140] Akanbi OA, Harms D, Wang B, Opaleye OO, Adesina O, Osundare FA, et al. Complete genome sequence of a hepatitis E virus genotype 1e strain from an outbreak in Nigeria, 2017. *Microbiol Resour Announc* 2019;8. <https://doi.org/10.1128/MRA.01378-18>. eCollection 2019 Jan.
- [141] Alfonsi V, Romano L, Ciccaglione AR, La Rosa G, Bruni R, Zanetti A, et al. Hepatitis E in Italy: 5 years of national epidemiological, virological and environmental surveillance, 2012 to 2016. *Euro Surveill* 2018;23. <https://doi.org/10.2807/1560-7917.ES.2018.23.41.1700517>.
- [142] Bruni R, Villano U, Equestre M, Chionne P, Madonna E, Trandeva-Bankova D, et al. Hepatitis E virus genotypes and subgenotypes causing acute hepatitis, Bulgaria, 2013–2015. *PLoS One* 2018;13:e0198045.
- [143] Di Pasquale S, De Santis P, La Rosa G, Di Domenico K, Iaconelli M, Micarelli G, et al. Quantification and genetic diversity of Hepatitis E virus in wild boar (*Sus scrofa*) hunted for domestic consumption in Central Italy. *Food Microbiol* 2019;82:194–201.
- [144] Jemersic L, Prpic J, Brnic D, Keros T, Pandak N, Dakovic Rode O. Genetic diversity of hepatitis E virus (HEV) strains derived from humans, swine and wild boars in Croatia from 2010 to 2017. *BMC Infect Dis* 2019;19: 269-019-3906-6.
- [145] Mirazo SD, Albora C, Quintero Gil D, Cabrera K, Ramos N, Orduz S, et al. A case of incidental infection of Hepatitis E virus (HEV) genotype 1 in a domestic pig. *Arch Virol* 2018;163:3403–7.
- [146] Realpe-Quintero M, Mirazo S, Viera-Segura O, Copado-Villagrana ED, Panduro A, Roman S, et al. Hepatitis E virus genotype 1 and hepatitis A virus dual infection in pediatric patients with a low socioeconomic status from Mexico. *Intervirology* 2018;61:105–10.
- [147] Suin V, Klamer SE, Hutse V, Wautier M, Jacques M, Abady M, et al. Epidemiology and genotype 3 subtype dynamics of hepatitis E virus in Belgium, 2010 to 2017. *Euro Surveill* 2019;24. <https://doi.org/10.2807/1560-7917.ES.2019.24.10.1800141>.
- [148] Viera-Segura O, Realpe-Quintero M, Panduro A, Roman S, Jose-Abrego A, Gonzalez-Aldaco K, et al. First detection of hepatitis E virus genotype 3 as a common infectious agent in patients with chronic liver damage in Mexico. *Ann Hepatol* 2019.
- [149] Wang B, Akanbi OA, Harms D, Adesina O, Osundare FA, Naidoo D, et al. A new hepatitis E virus genotype 2 strain identified from an outbreak in Nigeria, 2017. *Virology* 2018;15: 163-018-1082-8.
- [150] Boon D, Redd AD, Laeyendecker O, Engle RE, Nguyen H, Ocamo P, et al. Hepatitis E virus seroprevalence and correlates of anti-HEV IgG antibodies in the Rakai District, Uganda. *J Infect Dis* 2018;217:785–9.
- [151] Kuznetsova TV, Ivanova-Pozdejeva A, Reshetnjak I, Geller J, Varv K, Rumvolt R, et al. Hepatitis E virus infection in different groups of Estonian patients and people who inject drugs. *J Clin Virol* 2018;104:5–10.