VIRUS EVOLUTION

Virus Evolution, 2022, **8(1)**, 1–7
DOI: https://doi.org/10.1093/ve/veac028
Advance access publication date: 8 April 2022
Reflections

Reconstruction of the origin and dispersal of the worldwide dominant Hepatitis B Virus subgenotype D1

Nídia Sequeira Trovão,[1,*,†] Marijn Thijssen,[1] Bram Vrancken,[1,‡] Andrea-Clemencia Pineda-Peña[2,3] Thomas Mina,[4] Samad Amini-Bavil-Olyaee,[5] Philippe Lemey,[1,§] Guy Baele,[1,**] and Mahmoud Reza Pourkarim[1,6,7,*,††]

[1]Department of Microbiology, Immunology and Transplantation, KU Leuven, Rega Institute, Laboratory for Clinical and Epidemiological Virology, Herestraat 49, Leuven BE-3000, Belgium, [2]Global Health and Tropical Medicine, GHTM, Instituto de Higiene e Medicina Tropical, IHMT; Universidade Nova de Lisboa, UNL, Portugal Rua da Junqueira No 100, Lisbon 1349-008, Portugal, [3]Molecular Biology and Immunology Department, Fundación Instituto de Inmunología de Colombia (FIDIC); Faculty of Animal Science, Universidad de Ciencias Aplicadas y Ambientales (U.D.C.A.), Avenida 50 No. 26-20, Bogotá 0609, Colombia, [4]Nonis Lab Microbiology - Virology Unit, Gregori Afxentiou 5, Limassol 4003, Cyprus, [5]Cellular Sciences Department, Biosafety Development Group, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, USA, [6]Health Policy Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran and [7]Blood Transfusion Research Centre, High Institute for Research and Education in Transfusion Medicine, Hemmat Exp. Way, Tehran 14665-1157, Iran
[†]https://orcid.org/0000-0002-2106-1166
[‡]https://orcid.org/0000-0001-6547-5283
[§]https://orcid.org/0000-0003-2826-5353
[**]https://orcid.org/0000-0002-1915-7732
[††]https://orcid.org/0000-0003-4620-6453
*Corresponding authors: E-mail: nidistrovao@gmail.com; mahmoudreza.pourkarim@kuleuven.be

Abstract

Hepatitis B is a potentially life-threatening liver infection caused by the hepatitis B virus (HBV). HBV-D1 is the dominant subgenotype in the Mediterranean basin, Eastern Europe, and Asia. However, little is currently known about its evolutionary history and spatio-temporal dynamics. We use Bayesian phylodynamic inference to investigate the temporal history of HBV-D1, for which we calibrate the molecular clock using ancient sequences, and reconstruct the viral global spatial dynamics based, for the first time, on full-length publicly available HBV-D1 genomes from a wide range of sampling dates. We pinpoint the origin of HBV subgenotype D1 before the current era (BCE) in Turkey/Anatolia. The spatial reconstructions reveal global viral transmission with a high degree of mixing. By combining modern-day and ancient sequences, we ensure sufficient temporal signal in HBV-D1 data to enable Bayesian phylodynamic inference using a molecular clock for time calibration. Our results shed light on the worldwide HBV-D1 epidemics and suggest that this originally Middle Eastern virus significantly affects more distant countries, such as those in mainland Europe.

Key words: HBV; D1; Bayesian inference; MCMC; phylodynamics; full genome; temporal signal.

1. Introduction

In an effort to reach the goals of the WHO viral hepatitis elimination programme, countries continue to rely on efficient treatment of hepatitis C virus and vaccine availability for hepatitis B virus (HBV) infection (Pourkarim et al. 2018). In spite of large-scale vaccination programmes implemented in many countries, an estimated 260 million people are still chronically infected with HBV (Pourkarim and Van Ranst 2011), and about 880,000 people die of its consequences every year (Lozano et al. 2012; Pourkarim et al. 2014a). Genotype D is globally the most prevalent among the HBV genotypes (Pourkarim et al. 2010b; Paraskevis et al. 2013; Pourkarim et al. 2014b), and its 12 subgenotypes (D1–D12) have varying ethno-geographic dispersal patterns (Pineda-Pena et al. 2015: Hundie et al. 2016; Thijssen et al. 2020). HBV subgenotype D1 is most prevalent in Asia and in the Mediterranean basin and is also found in Europe, especially in Eastern European countries. The D2 and D3 subgenotypes are most often found in Europe and East Asia, whereas the subgenotypes D4–D12 show a more restricted spread in particular among ethnic groups belonging to different continents (Hundie et al. 2016; Banerjee et al. 2014; Thijssen et al. 2020).

The natural history of HBV infection is associated with a diverse and broad spectrum of outcomes from acute, inactive to active chronic infection, which may progress to liver cirrhosis or hepatocellular carcinoma (Trepo, Chan, and Lok 2014; Mina et al. 2015). Roughly 43 per cent of HBV-infected people live in regions of intermediate prevalence (2–7 per cent), including south-central and southwest Asia, eastern and southern Europe, Russia, and central and south America. In these regions, mixed patterns of transmission are observed, including infant, childhood, and adult transmission (Trepo, Chan, and Lok 2014; Pourkarim et al. 2014a). Most of the infected patients in these regions are asymptomatic chronic carriers. It is estimated that between 40 and 80 per cent of these chronic carriers are not aware of their own infection status (Hahné et al. 2013), with different HBV genotypes

and subgenotypes presenting a variety of virological characteristics, such as different mutational patterns in different open reading frames (Grethe et al. 1998; Zaaijer, Bouter, and Boot 2008; Zehender et al. 2008).

Several studies have proposed differing explanations for the origin of HBV, including cospeciation in primates, coevolution following the major migration of the anatomically modern human, cross-species transmission between humans and non-humans, or a bat origin of the virus (Littlejohn, Locarnini, and Yuen 2016). Previous studies have also sought to estimate the origins of HBV genotype D using maximum-likelihood methods, pointing its putative origin to be in North Africa or the Middle East (Kostaki et al. 2018).

The burgeoning field of pathogen phylodynamics research has mainly focused on fast(er) evolving pathogens, particularly RNA viruses affecting a plethora of hosts (Trovao et al. 2015a,b; Vrancken et al. 2014), with comparatively fewer analyses targeting double-stranded DNA viruses such as HBV. However, HBV uses an RNA intermediate called pre-genomic RNA and reverse transcriptase for its genome replication and is therefore expected to evolve faster than other DNA viruses. It is unclear to what extent phylodynamic concepts apply to HBV-D1 because their ancestral history may be masked by their slower evolutionary rate and consequently weaker temporal signal (Trovao et al. 2015a). This hypothesis was recently corroborated by a study reporting paradoxical evolutionary patterns for HBV and suggesting that molecular clock-based methods may have limited utility, particularly for very long timeframes, with the risk of underestimating the time to the most recent common ancestor (MRCA) (Ross et al. 2018). On the other hand, a recent study has shown the usefulness of ancient HBV sequences in securing sufficient temporal signals to calibrate the molecular clock (Hahné et al. 2013). The goal of this study is to investigate the spatial processes that gave rise to the current HBV-D1 epidemic using a Bayesian phylodynamics framework that appropriately characterizes the uncertainty associated with the estimation process. Phylodynamic methods enable estimating the tempo and time scale of HBV-D1 evolution and inferring the spatial population structure, shedding light on the spatio-temporal patterns that determined the current geographic distribution of this subgenotype while unveiling the migration links that connect the global HBV-D1 epidemics. Here, we use a comprehensive collection of complete genome sequences to fill this gap and help inform and implement surveillance in the framework of public health policies.

2. Results

We generated comprehensive genetic datasets and investigated their temporal signal followed by comprehensive Bayesian phylogenetic and phylogeographic analyses to estimate HBV-D1 timed evolutionary histories and reconstruct the viral spatial diffusion (see detailed description in Supplementary File 1).

2.1 Sequence evolution and temporal analysis

We estimated the evolutionary rate for the reference dataset B to be 2.17×10^{-5} (95 per cent highest posterior density (HPD): 1.25×10^{-5}–3.26×10^{-5}) substitutions per site per year, and the corresponding time to the most recent common ancestor (tMRCA) for HBV genotype D at 1617 BCE (95 per cent HPD: 3320 BCE—421 BCE) (Supplementary Figure S1), which is in line with estimates from a recent study that includes additional ancient samples collected from Eurasians and Native Americans (Kocher et al. 2021). We used this evolutionary rate estimate as an informative normal prior distribution on the mean evolutionary rate (mean = 2.17×10^{-5}; standard deviation = 5.14×10^{-6}) for subsequent reconstruction of dataset A (full HBV-D1 dataset obtained from GenBank after exclusion of recombinant sequences) and dataset C (homogeneously distributed dataset that mitigates sampling bias for the location trait while maximizing the capture of the phylogenetic diversity). The mean evolutionary rate estimated for dataset A was estimated at 2.71×10^{-5} (95 per cent highest posterior density (HPD): 1.80×10^{-5}–3.78×10^{-5}) substitutions per site per year and the corresponding tMRCA for HBV-D1 at 1308 BCE (95 per cent HPD: 2848 BCE—166 BCE). The viral population dynamics appear constant over time. A similar demographic pattern was also observed for dataset C, for which we estimated an evolutionary rate of 2.72×10^{-5} (95 per cent HPD: 1.79×10^{-5}–3.82×10^{-5}) substitutions/site/year and a tMRCA of 1237 BCE (95 per cent HPD: 2816 BCE—130 BCE).

Table 1. Association index to study phylogeographic structure.

Dataset	Number of sequences	Sampling criteria	Association index
A	643	All sequences	0.23 (95% HPD: 0.21–0.25)
C	583	Maximum phylogenetic diversity	0.23 (95% HPD: 0.21–0.25)

As expected from using an informative rate prior based on dataset B when analysing datasets A and C, we estimated similar posterior evolutionary rate distributions across datasets. The somewhat older tMRCA estimated for dataset B compared to those of datasets A and C might be due to dataset B being comprised of sequences of different subgenotypes within genotype D, whereas datasets A and C only have sequences from genotype D1.

2.2 Discrete phylogeography

To assess the phylogeographic structure in the datasets, we quantified the degree of phylogenetic clustering by trait summarized using the association index (AI) over all nodes in the MCC tree. Table 1 lists AI metrics for datasets A and C, pointing to a high degree of phylogeny-trait association by location and the capability to inform the discrete diffusion analyses.

We reconstructed the spatial diffusion dynamics underlying the HBV-D1 global spread for both datasets. Datasets A and C provided strong support (posterior probabilities of 0.89 and 0.97 for dataset A and C, respectively) for an origin of the current HBV-D1 diversity in Turkey or the ancient region of Anatolia (Figs 1 and 2, Tables 2 and 3). The spatial estimates, therefore, trace the HBV-D1 origin and maintenance to Turkey, after which it first spread to Iran, and later on to Europe, Asia, Syria, Africa, and India. HBV-D1 reached New Zealand via an introduction from India. The period between the late 1400s and 1800s was also marked by a high degree of viral interchange between the Middle East, Asia, and Europe, consistent with the events that took place during the Early Modern Period, including the Ages of Discovery and Enlightenment (Movies 1 and 2 for datasets A and C, respectively).

Apart from summarizing the spatio-temporal history of HBV-D1, we also quantified the support for different diffusion pathways as Bayes factor (BF) support for non-zero rates of discrete location transitions between all pairs of 27 countries and 9 regions for datasets A and C, respectively, in order to identify those well-supported viral transmissions that best explain the observed current spread. We found that the movements from Turkey to Iran, Syria, India, and Europe, as well as Iran to Turkey for dataset A (Fig. 1 and Table 2), and the movements from Turkey to Europe and Iran to Turkey for dataset C (Fig. 2 and Table 3) are supported

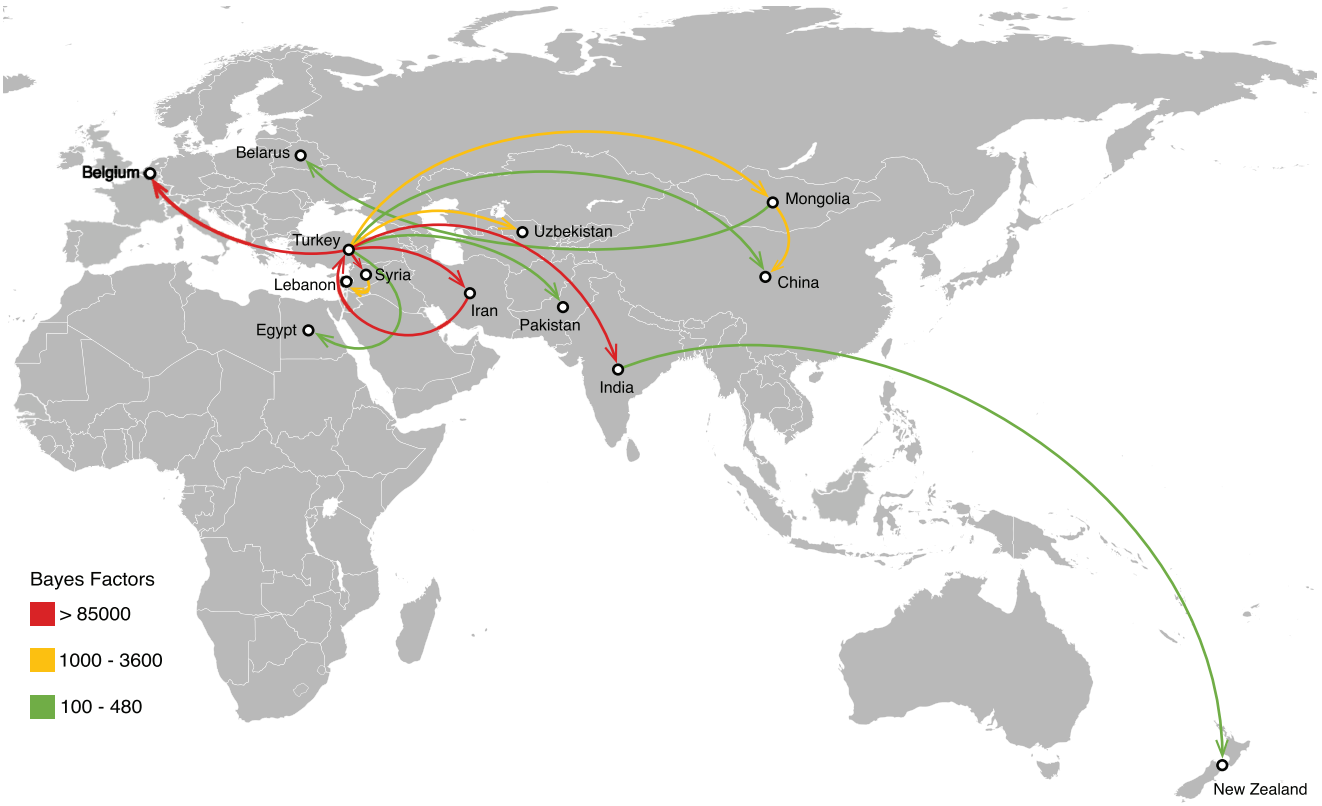

Figure 1. Bayes factor test support for discrete diffusion rates inferred for dataset A. Locations are represented by the centroid coordinate of the country/region. Rates supported by a BF >80,000 are indicated.

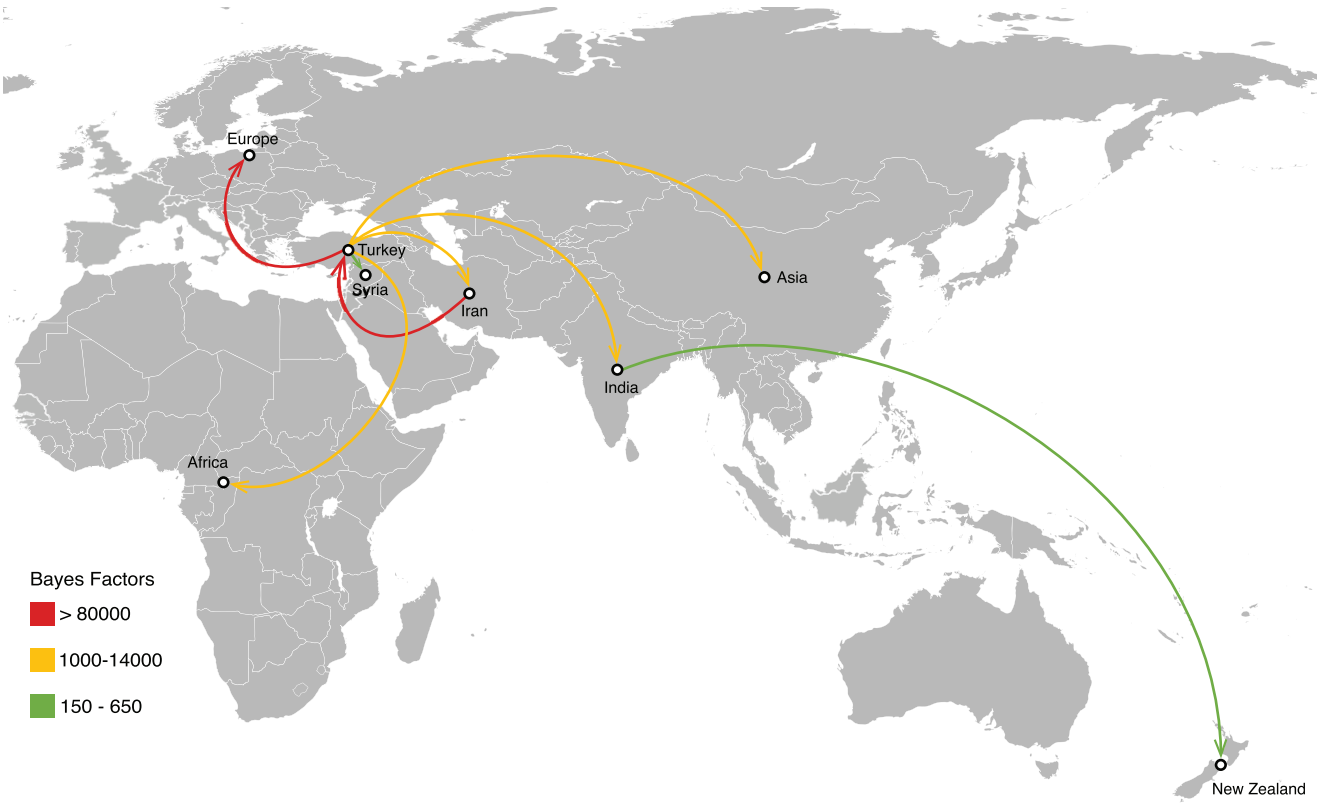

Figure 2. Bayes factor test support for discrete diffusion rates inferred for dataset C. Locations are represented by the centroid coordinate of the country/region. Rates supported by a BF > 80,000 are indicated.

Table 2. Asymmetrical heat maps of HBV-D1 flow between locations for dataset A[a].

		To												
		Belarus	Belgium	China	Egypt	India	Iran	Lebanon	Mongolia	NZ	Pakistan	Syria	Turkey	Uzbekistan
From	India	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.38	0.01	0.00	0.00	0.00
	Iran	0.00	0.09	0.00	0.00	0.24	0.00	0.22	0.00	0.00	0.00	1.18	5.81	0.04
	Mongolia	0.75	0.00	2.56	0.00	0.01	0.01	0.00	0.00	0.00	0.04	0.00	0.01	0.02
	Syria	0.01	0.41	0.00	0.16	0.00	0.00	9.70	0.00	0.00	0.00	0.00	0.07	0.00
	Turkey	0.48	16.76	1.89	1.25	6.09	17.91	2.13	3.42	0.03	1.53	12.21	0.00	2.65

NZ: New Zealand
[a]Markov jump counts measure the expected number of viral movements that occur along the branches of the phylogeny, providing a measure of gene flow. The intensity of the color (red = high; green = low) reflects the percentage of Markov jump counts from location of origin (y-axis) to a destination (x-axis). Transitions shown are supported by BF >100.

Table 3. Asymmetrical heat maps of HBV-D1 flow between locations for dataset D[a].

		To							
		Africa	Asia	Europe	India	Iran	NZ	Syria	Turkey
From	India	0.01	0.01	0.03	0.00	0.03	0.52	0.02	0.01
	Iran	0.00	0.11	0.07	0.22	0.00	0.00	0.97	5.62
	Turkey	3.69	15.22	25.54	6.74	17.17	0.02	10.27	0.00

NZ: New Zealand
[a]Markov jump counts measure the expected number of viral movements that occur along the branches of the phylogeny, providing a measure of gene flow. The intensity of the color (dark blue = high; white = low) reflects the percentage of Markov jump counts from location of origin (y-axis) to a destination (x-axis). Transitions shown are supported by BF >100.

by very high BF values (>80,000). We complemented the support for these rates by estimating the number of transitions that occurred between the states involved using Markov jump counts for transitions supported by BF values >100 (Kass and Raftery 1995) (Tables 2 and 3). These estimates indicate that the location that seeds most HBV-D1 to other regions is Turkey/Anatolia, with viral jumps to Europe encompassing more than 25 per cent of all mean Markov jump counts (Table 3). The second most frequent viral transition in both datasets occurs from Turkey to Iran, encompassing approximately 17 per cent of all Markov jump counts (Tables 2 and 3).

3. Discussion

HBV Genotype D is the dominant genotype worldwide. Its subgenotype D1 is the most prevalent strain in south-east Europe, the Mediterranean Basin, and the south, south-west, and central regions of Asia (Zehender et al. 2012a), (Pourkarim et al. 2009), (Schweitzer et al. 2015). In 2009, the WHO estimated that around 4.3 million persons are infected with HBV in the Mediterranean region (Melhem et al. 2015). Viral transmission of HBV occurs horizontally or perinatally in childhood and by infected body fluids during adulthood. The routes of viral transmission are influenced by socio-economic status, lifestyle, occupation, and cultural practices (Alavian et al. 2008; Zidan et al. 2012). Although mass vaccination programmes against HBV infection were crucial in the dramatic decrease of the chronicity rate worldwide (Zahedi et al. 2012), there are still thousands of new cases of HBV each year (Trepo, Chan, and Lok 2014; Graber-Stiehl 2018; Thijssen et al. 2019; Wiktor, Borawski, and Stępień 2019).

In this study, we used a comprehensive dataset of 648 full-length genome sequences to estimate the global evolutionary dynamics of HBV-D1, which included five ancient sequences. To our knowledge, we present the largest dataset of full-length HBV-D1 sequences analysed so far. It has been demonstrated that only full-length genomes of HBV allow for sufficient phylogenetic resolution (Pourkarim et al. 2009; Schaefer, Magnius, and Norder 2009; Pourkarim et al. 2010a,b; 2011; Amini-Bavil-Olyaee et al. 2012; Pineda-Pena et al. 2015) and using partial genomes can mislead evolutionary studies (Olinger et al. 2006; Hubschen et al. 2011; Vrancken et al. 2018).

Over the past half century, HBV-D1 has constituted a burden to low endemicity countries receiving immigrants from intermediate to high endemicity countries (Coppola et al. 2015; Thijssen et al. 2019). Control strategies and prophylaxis would benefit from a better understanding of the HBV-D1 epidemic dynamics. Specifically, public health programmes could redirect efforts by targeting immigrant populations from intermediate to high endemicity countries to prevent viral transmission (Thijssen et al. 2019).

The genomic structure of HBV represents a challenge for reconstructions of the spatio-temporal history. Despite replicating through an error-prone reverse transcriptase, HBV has a partially double-stranded DNA genome encompassing four partially overlapping genes. This provides a conserved genetic population and evolutionary constraint requiring the use of three independent codon partitions. Additionally, it hinders the accumulation of temporal signal, despite the large temporal spread in available sampling dates (see Supplementary Figure S2). Because of the lack of temporal signal, we used five ancient sequences in order to calibrate the molecular clock model and estimate time-measured phylogenetic reconstructions. We also used the evolutionary rate derived from temporal analysis of dataset B as an informative normal prior distribution on the mean evolutionary rate (mean = 2.17×10^{-5}; standard deviation = 5.14×10^{-6}).

We estimated that HBV-D1 is evolving at an average rate of 2.71×10^{-5} substitutions per site per year with very limited variation throughout its history. This result is almost 16 (95 per cent HPD: 0.17–17.46) times slower than the only previous estimate of HBV-D1 evolutionary rate derived from 15 HBsAg/reverse transcriptase fragment sequences from more recent isolates (Ciccozzi et al. 2013). Nonetheless, the resulting evolutionary rate and tMRCA estimate, 1308 BCE (95 per cent HPD: 2848 BCE–166 BCE), are in line with recent estimates that use the ancient samples to calibrate the molecular clock (Zehender et al. 2012a,b; Ciccozzi

et al. 2013; Mühlemann et al. 2018). This illustrates the need for comprehensive datasets and appropriate model specifications that enable a more accurate reconstruction of the viral evolutionary history.

Because of the high level of sampling heterogeneity across countries in our datasets, with most sequences originating from Belgium ($n = 52$), Iran ($n = 175$), India ($n = 84$), Syria ($n = 64$), and Turkey ($n = 93$) (Supplementary Figure S3), we created a more balanced dataset C to alleviate the potential biasing impact of uneven sampling (Supplementary Figure S4) for evaluating the HBV-D1 spatio-temporal dynamics using a discrete phylogeographic approach (Lemey, Suchard, and Rambaut 2009). This fact that we identify Turkey, corresponding to the ancient region of Anatolia, as the geographical origin from where HBV-D1 spread worldwide in both datasets, in line with previous studies (Paraskevis et al. 2013; Kostaki et al. 2018), provides reassurance that this finding is not significantly impacted by sampling bias.

Multiple introductions of HBV-D1 were estimated to have occurred in the countries and regions included in this study, consistent with results from previous studies (Banerjee et al. 2006; Pourkarim et al. 2010b; Amini-Bavil-Olyaee et al. 2011; Zehender et al. 2012a; Zehender et al. 2014) and the global distribution of this subgenotype (Pourkarim et al. 2008, 2014a). We also estimated the highest number of viral transitions from Turkey to several other locations, which were highly supported by BF values.

Future spatial analyses of HBV may also benefit from structured coalescent approaches that appear to be less sensitive to sampling bias. Their current computationally demands limit application to large datasets (Vaughan et al. 2014), but promising approximate algorithms have recently been developed (De Maio et al. 2015). By subsampling the dataset using two different approaches (randomized and according to maximum phylogenetic divergence), we were able to produce consistent results for the temporal and phylogeographic reconstructions. The lack of dense sampling of the HBV-D1 epidemics, socio-demographic information and routes of transmission limits our endeavors to investigate the social transmission trends of HBV-D1. This represents a particular limitation for regions with high HBV-D1 prevalence but disproportionately low number of samples such as the region of the Americas (2–7 per cent of the population is infected, but less than 20 sequences from this region are available) (MacLachlan and Cowie 2015). Nonetheless, our spatio-temporal inferences uncovered that the majority of transmission events that gave rise to HBV-D1 spread worldwide occurred in the periods encompassing the Early Modern Period, including the Ages of Discovery and Enlightenment. Population sizes and densities, and human movement networks are additional factors that could help unveil the underlying processes that shaped the dynamics of viral spread if investigated using the generalized linear model extension of the discrete phylogeographic model (Lemey et al. 2014). The complexity of reconstructing HBV evolutionary history might be driven by various factors, including transmission bottlenecks (Lin et al. 2016; Du et al. 2017), differences between inter-/intra-host diversity (Ramachandran et al. 2011), different evolutionary rates in different body compartments (Datta et al. 2009; Sinha et al. 2019), and long duration of infection with a stable archived reservoir of cccDNA (Vrancken, Suchard, and Lemey 2017; Li, Sohn, and Seeger 2018; Brezgin et al. 2019). We also hypothesized that the challenge in reaching convergence for the temporal parameters, particularly the date of divergence and coalescent and evolutionary rate, was due to the very constrained genomic structure with overlapping reading frames, varying mutation and replication rates, as well as selective pressures (Trovao et al. 2015a). However, Patterson Ross et al. (2018) observed a lack of temporal signal for HBV-D3 even when comparing old (16th century) and modern sequences. This implies that temporal reconstruction of HBV should be interpreted cautiously, and whereas the viral diffusion pathways were highly supported, its temporal history is highly dependent on appropriate rate calibration.

4. Conclusion

In conclusion, we have used a comprehensive phylogeographic approach to determine the worldwide evolutionary and spatio-temporal dynamics of HBV-D1. This work exalts the importance of careful methodological calibration to unveil the origin of epidemics and the migration pathways that the virus explored for its expansion while accommodating its deep evolutionary history. We hope that our findings may ultimately help inform public health measures towards control and strategies for increased preparedness.

Data availability

The statistical code and datasets are available at https://github.com/nidiatrovao/HBVD1-Trovao-et-al.

Supplementary data

Supplementary data is available at *Virus Evolution* online.

Funding

N.S.T. and P.L. were supported by the European Union Seventh Framework Programme [FP7/2007-2013] under Grant Agreement number 278433-PREDEMICS. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 725422 - ReservoirDOCS). MT is a PhD fellow at the Research Foundation Flanders (FWO, Belgium, grant number 1S47118N). A.-C.P.-P. was supported by European Funds through grant 'Bio-Molecular and Epidemiological Surveillance of HIV Transmitted Drug Resistance, Hepatitis Co-Infections and Ongoing Transmission Patterns in Europe' (BEST HOPE) (project funded through HIVERA: Harmonizing Integrating Vitalizing European Research on HIV/Aids, grant 249697); by Fundação para a Ciência e Tecnologia for funds to GHTM-UID/Multi/04413/2013; by the Migrant HIV project (financed by FCT: PTDC/DTP-EPI/7066/2014; and by Gilead Génese HIVLatePresenters. B.V. was supported by a postdoctoral grant (12U7121N) of the FWO (Fonds Wetenschappelijk Onderzoek – Vlaanderen). G.B. acknowledges support from the Interne Fondsen KU Leuven/Internal Funds KU Leuven under grant agreement C14/18/094 and the Research Foundation – Flanders ('Fonds voor Wetenschappelijk Onderzoek – Vlaanderen', G0E1420N, G098321N). This work was supported by the Bijzonder Onderzoeksfonds KU Leuven (BOF) No. OT/14/115. This work was supported by public grants. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest: The authors declare that they have no competing interests.

Author contributions

The study was conceived and designed by N.S.T, M.T., A.C.P.P., P.L. S.A.B.O., and M.R.P. A.C.P.P. and T.M. collected the data. N.S.T. assembled the datasets. The data were analysed and interpreted by N.S.T., G.B., and M.R.P. N.S.T. drafted the article. N.S.T., B.V., G.B., P.L., S.A.B.O., and M.R.P. critically revised the article. B.V., G.B., P.L., and M.R.P. obtained funding. All authors read and approved the final manuscript.

References

Alavian, S. M. et al. (2008) 'Hepatitis B Virus Infection in Iran: A Systematic Review', *Hepatitis Monthly*, 8: 281–94.

Amini-Bavil-Olyaee, S. et al. (2012) 'Providing Strong Evidence of Nosocomial Outbreak of Hepatitis B Virus Infection', *The Journal of Hospital Infection*, 80: 269–70. author reply 70–2.

Amini-Bavil-Olyaee, S. et al. (2011) 'Single-step Real-time PCR to Quantify Hepatitis B Virus and Distinguish Genotype D from non-D Genotypes', *Journal of Viral Hepatitis*, 18: 300–4.

Banerjee, A. et al. (2006) 'Phylogenetic Relatedness and Genetic Diversity of Hepatitis B Virus Isolates in Eastern India', *Journal of Medical Virology*, 78: 1164–74.

Banerjee, P. et al. (2014) 'A Rare HBV Subgenotype D4 with Unique Genomic Signatures Identified in North-eastern India—an Emerging Clinical Challenge?' *PLoS One*, 9: e109425.

Brezgin, S. et al. (2019) 'Replenishment of Hepatitis B Virus cccDNA Pool Is Restricted by Baseline Expression of Host Restriction Factors in Vitro', *Microorganisms*, 7: 533.

Ciccozzi, M. et al. (2013) 'Molecular Analysis of Hepatitis B Virus in Bulgaria', *Journal of Medical Virology*, 85: 49–54.

Coppola, N. et al. (2015) 'Hepatitis B Virus Infection in Immigrant Populations', *World Journal Hepatology*, 7: 2955–61.

Datta, S. et al. (2009) 'Genetic Characterization of Hepatitis B Virus in Peripheral Blood Leukocytes: Evidence for Selection and Compartmentalization of Viral Variants with the Immune Escape G145R Mutation', *Journal of Virology*, 83: 9983–92.

De Maio, N. et al. (2015) 'New Routes to Phylogeography: A Bayesian Structured Coalescent Approximation', *PLoS Genetics*, 11: e1005421.

Du, Y. et al. (2017) 'Quantifying Perinatal Transmission of Hepatitis B Viral Quasispecies by Tag Linkage Deep Sequencing', *Scientific Reports*, 7: 1–11.

Graber-Stiehl, I. (2018) 'The Silent Epidemic Killing More People than HIV, Malaria or TB', *Nature*, 564: 24–7.

Grethe, S. et al. (1998) 'Characterization of Unusual Escape Variants of Hepatitis B Virus Isolated from a Hepatitis B Surface Antigen-negative Subject', *Journal of Virology*, 72: 7692–6.

Hahné, S. J. et al. (2013) 'Infection with Hepatitis B and C Virus in Europe: A Systematic Review of Prevalence and Cost-effectiveness of Screening', *BMC Infectious Diseases*, 13: 181.

Hubschen, J. M. et al. (2011) 'Detection of A New Subgenotype of Hepatitis B Virus Genotype A in Cameroon but Not in Neighbouring Nigeria', *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 17(1): 88–94.

Hundie, G. B. et al. (2016) 'A Novel Hepatitis B Virus Subgenotype D10 Circulating in Ethiopia', *Journal of Viral Hepatitis*, Feb 24(2): 163–73.

Kass, R. E., and Raftery, A. E. (1995) 'Bayes Factors', *Journal of the American Statistical Association*, 90: 773–95.

Kocher, A. et al. (2021) 'Ten Millennia of Hepatitis B Virus Evolution', *Science (New York, NY)*, 374: 182–8.

Kostaki, E.-G. et al. (2018) 'Unravelling the History of Hepatitis B Virus Genotypes A and D Infection Using A Full-genome Phylogenetic and Phylogeographic Approach', *eLife*, 7: e36709.

Lemey, P. et al. (2014) 'Unifying Viral Genetics and Human Transportation Data to Predict the Global Transmission Dynamics of Human Influenza H3N2', *PLoS Pathogens*, 10: e1003932.

Lemey, P., Suchard, M., and Rambaut, A. (2009) 'Reconstructing the Initial Global Spread of a Human Influenza Pandemic: A Bayesian Spatial-temporal Model for the Global Spread of H1N1pdm', *PLoS Currents*, 1: RRN1031.

Li, M., Sohn, J. A., and Seeger, C. (2018) 'Distribution of Hepatitis B Virus Nuclear DNA', *Journal of Virology*, 92: e01391–17.

Lin, S. Y. et al. (2016) 'Molecular Evolution and Phylodynamics of Acute Hepatitis B Virus in Japan', *PLoS One*, 11: 1–14.

Littlejohn, M., Locarnini, S., and Yuen, L. (2016) 'Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus', *Cold Spring Harbor Perspectives in Medicine*, 6: a021360.

Lozano, R. et al. (2012) 'Global and Regional Mortality from 235 Causes of Death for 20 Age Groups in 1990 and 2010: A Systematic Analysis for the Global Burden of Disease Study 2010', *The Lancet*, 380: 2095–128.

MacLachlan, J. H., and Cowie, B. C. (2015) 'Hepatitis B Virus Epidemiology', *Cold Spring Harbor Perspectives in Medicine*, 5: a021410.

Melhem, N. M. et al. (2015) 'Human Immunodeficiency Virus and Viral Hepatitis among High-risk Groups: Understanding the Knowledge Gap in the Middle East and North Africa Region', *World Journal of Hepatology*, 7: 2619–30.

Mina, T. et al. (2015) 'Genomic Diversity of Hepatitis B Virus Infection Associated with Fulminant Hepatitis B Development', *Hepatitis Monthly*, 15: e29477.

Mühlemann, B. et al. (2018) 'Ancient Hepatitis B Viruses from the Bronze Age to the Medieval Period', *Nature*, 557: 418–23.

Olinger, C. M. et al. (2006) 'Phylogenetic Analysis of the Precore/core Gene of Hepatitis B Virus Genotypes E and A in West Africa: New Subtypes, Mixed Infections and Recombinations', *The Journal of General Virology*, 87: 1163–73.

Paraskevis, D. et al. (2013) 'Dating the Origin and Dispersal of Hepatitis B Virus Infection in Humans and Primates', *Hepatology (Baltimore, Md)*, 57: 908–16.

Pineda-Pena, A. C. et al. (2015) 'Epidemiological History and Genomic Characterization of non-D1 HBV Strains Identified in Iran', *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, 63: 38–41.

Pourkarim, M. R. et al. (2014a) 'Molecular Identification of Hepatitis B Virus Genotypes/subgenotypes: Revised Classification Hurdles and Updated Resolutions', *World Journal of Gastroenterology*, 20: 7152–68.

Pourkarim, M. R. et al. (2010a) 'Are Hepatitis B Virus "subgenotypes" Defined Accurately?' *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, 47: 356–60.

Pourkarim, M. R. et al. (2011) 'HBV Subgenotype Misclassification Expands Quasi-subgenotype A3', *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 17: 947–9.

Pourkarim, M. R. et al. (2010b) 'Molecular Evolutionary Analysis and Mutational Pattern of Full-length Genomes of Hepatitis B Virus Isolated from Belgian Patients with Different Clinical Manifestations', *Journal of Medical Virology*, 82: 379–89.

Pourkarim, M. R. et al. (2010c) 'Novel Hepatitis B Virus Subgenotype A6 in African-Belgian Patients', *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, 47: 93–6.

Pourkarim, M. R. et al. (2018) 'Iran's Hepatitis Elimination Programme Is under Threat', *The Lancet*, 392: 1009.

Pourkarim, M. R., and Van Ranst, M. (2011) 'Guidelines for the Detection of a Common Source of Hepatitis B Virus Infections', *Hepatitis Monthly*, 11: 783–5.

Pourkarim, M. R. et al. (2009) 'Phylogenetic Analysis of Hepatitis B Virus Full-length Genomes Reveals Evidence for a Large Nosocomial Outbreak in Belgium', *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, 46: 61–8.

Pourkarim, M. R. et al. (2014b) 'Molecular Characterization of Hepatitis B Virus (HBV) Strains Circulating in the Northern Coast of the Persian Gulf and Its Comparison with Worldwide Distribution of HBV Subgenotype D1', *Journal of Medical Virology*, 86: 745–57.

Pourkarim, M. R. et al. (2008) 'An Aberrant High Prevalence of Hepatitis B Infection among Afghans Residing in One of the Bushehr Refugee Camps (Dalaki Camp) in the Southwest of Iran', *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases*, 12: 101–2.

Ramachandran, S. et al. (2011) 'Evaluation of Intra-host Variants of the Entire Hepatitis B Virus Genome', *PLoS One*, 6(9): e25232, 1–10.

Ross, Z. P. et al. (2018) 'The Paradox of HBV Evolution as Revealed from a 16th Century Mummy', *PLoS Pathogens*, 14: e1006750.

Schaefer, S., Magnius, L., and Norder, H. (2009) 'Under Construction: Classification of Hepatitis B Virus Genotypes and Subgenotypes', *Intervirology*, 52: 323–5.

Schweitzer, A. et al. (2015) 'Estimations of Worldwide Prevalence of Chronic Hepatitis B Virus Infection: A Systematic Review of Data Published between 1965 and 2013', *The Lancet*, Oct 17; 386(10003): 1546–55.

Sinha, M. et al. (2019) 'Pro-oncogenic, Intra Host Viral Quasispecies in Diffuse Large B Cell Lymphoma Patients with Occult Hepatitis B Virus Infection', *Scientific Reports*, 9(1): 14516, 1–8.

Thijssen, M. et al. (2019) 'Mass Migration to Europe: An Opportunity for Elimination of Hepatitis B Virus?' *The Lancet Gastroenterology & Hepatology*, 4: 315–23.

Thijssen, M. et al. (2020) 'Novel Hepatitis B Virus Subgenotype A8 and Quasi-subgenotype D12 in African–Belgian Chronic Carriers', *International Journal of Infectious Diseases*, 93: 98–101.

Trepo, C., Chan, H. L., and Lok, A. (2014) 'Hepatitis B Virus Infection', *Lancet (London, England)*, 384: 2053–63.

Trovao, N. S. et al. (2015a) 'Host Ecology Determines the Dispersal Patterns of a Plant Virus', *Virus Evolution*, 1: vev016.

Trovao, N. S. et al. (2015b) 'Bayesian Inference Reveals Host-Specific Contributions to the Epidemic Expansion of Influenza A H5N1', *Molecular Biology Evolution*, 32: 3264–75.

Vaughan, T. G. et al. (2014) 'Efficient Bayesian Inference under the Structured Coalescent', *Bioinformatics*, 30: 2272–9.

Vrancken, B. et al. (2018) 'Why Comprehensive Datasets Matter When Inferring Epidemic Links or Subgenotyping', *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 65: 350–1.

Vrancken, B. et al. (2014) 'The Genealogical Population Dynamics of HIV-1 in a Large Transmission Chain: Bridging within and among Host Evolutionary Rates', *PLoS Computational Biology*, 10: e1003505.

Vrancken, B., Suchard, M. A., and Lemey, P. (2017) 'Accurate Quantification of Within- and Between-host HBV Evolutionary Rates Requires Explicit Transmission Chain Modelling', *Virus Evolution*, 3: vex028.

Wiktor, A., Borawski, T., and Stępień, M. (2019) 'Hepatitis B in Poland in 2017', *Przeglad Epidemiologiczny*, 73: 157–66.

Zaaijer, H. L., Bouter, S., and Boot, H. J. (2008) 'Substitution Rate of the Hepatitis B Virus Surface Gene', *Journal of Viral Hepatitis*, 15: 239–45.

Zahedi, M. J. et al. (2012) 'Seroprevalence of Hepatitis Viruses B, C, D and HIV Infection among Hemodialysis Patients in Kerman Province, South-East Iran', *Hepatitis Monthly*, 12: 339–43.

Zehender, G. et al. (2012a) 'Spatial and Temporal Dynamics of Hepatitis B Virus D Genotype in Europe and the Mediterranean Basin', *PLoS One*, 7: e37198.

Zehender, G. et al. (2014) 'Enigmatic Origin of Hepatitis B Virus: An Ancient Travelling Companion or a Recent Encounter?' *World Journal of Gastroenterology*, 20: 7622–34.

Zehender, G. et al. (2008) 'Different Evolutionary Rates and Epidemic Growth of Hepatitis B Virus Genotypes A and D', *Virology*, 380: 84–90.

Zehender, G. et al. (2012b) 'Reconstruction of the Epidemic History of Hepatitis B Virus Genotype D in Albania', *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 12: 291–8.

Zidan, A. et al. (2012) 'Epidemiological Pattern of Hepatitis B and Hepatitis C as Etiological Agents for Hepatocellular Carcinoma in Iran and Worldwide', *Hepatitis Monthly*, 12: e6894.