

RESEARCH ARTICLE

Open Access



# Is Hepatitis Delta infections important in Brazil?

Maira Ferreira Cicero<sup>1</sup>, Nathalia Mantovani Pena<sup>1</sup>, Luiz Claudio Santana<sup>1</sup>, Rafael Arnold<sup>1</sup>,  
Rafael Gonçalves Azevedo<sup>1</sup>, Élcio de Souza Leal<sup>2</sup>, Ricardo Sobhie Diaz<sup>1</sup> and Shirley Vasconcelos Komninakis<sup>1,3\*</sup>

## Abstract

**Background:** The Hepatitis Delta Virus (HDV) can increase the incidence of fulminant hepatitis. For this infection occurs, the host must also be infected with Hepatitis B Virus. Previous studies demonstrated the endemicity and near exclusivity of this infection in the Amazon region, and as a consequence of the difficulty in accessing this area we used dried blood spots (DBS) in sample collection. The aims of this study were to investigate the presence of recombination, to analyze the epidemiology, ancestry and evolutionary pressures on HDV in Brazil.

**Methods:** Blood samples from 50 individuals were collected using dried-blood spots (DBS 903, Whatman), and sent via regular mail to Retrovirology Laboratory from Federal University of São Paulo, where the samples were processed. In the analysis the following software were used: PhyML, RDP, BEAST, jModelTest and CODEML.

**Results:** Our results confirm the prevalence of HDV-3 in the Amazon region of Brazil, with the absence of inter-genotypic recombination. It was identified a positive selection in probable epitopes of HDV on B lymphocytes that might indicate that the virus is changing to escape the humoral response of the host. The analysis of the time of the most common ancestor demonstrated the exponential growth of this virus in late 1970s that lasted until 1995, after which it remained constant. It was also observed a probable founder effect in two cities, which demonstrate the need to focus on prevention methods against HBV/HDV infection.

**Conclusion:** We confirmed the prevalence of HDV-3 in the Amazon region of Brazil, without inter-genotypic recombination. The analysis of the time of the most common ancestor showed that this infection remain constant in the studied area. Taking into account the probable founder effect established in the cities of Rio Branco and Porto Velho, a focus on preventive methods is recommended against these infections.

**Keywords:** Hepatitis Delta virus, Dried-blood spot, Founder effect, Phylogeography, Phylogenetic analysis, tMRCA, Brazil

## Background

Hepatitis delta is a disease that has been neglected by the health system [1], despite being responsible for the infection of 15 to 20 million people worldwide [2, 3]. In Brazil, 77 % of the infections with Hepatitis Delta Virus (HDV) occur in the North region.

The Northern region of Brazil has locations that are geographically isolated, which hinders the samples transportation until the processing laboratory. Thereat, it is used the

dried blood spot (DBS) in sample collection at this region, for monitoring the HDV viral load (unpublished data). The use of DBS for sample collection and transportation is widely used in the HIV field, for monitoring the HIV viral load and the selection of drug resistance, on samples from remote sites or with limited financial resources. The results obtained using DBS sample collection are the same as using vacuum tubes [4–6].

HDV is the only subviral satellite agent known to infect humans. This virus requires the protection provided by the envelope of the Hepatitis B Virus (HBV) to enable its budding process from infected cells. Therefore, hepatitis delta occurs only in patients previously infected with HBV [7]. This co-infection increases the incidence

\* Correspondence: skomninakis@yahoo.com.br

<sup>1</sup>Retrovirology Laboratory, Federal University of São Paulo, 781 Pedro de Toledo Street, São Paulo, São Paulo, Brazil

<sup>3</sup>School of Medicine of ABC (FMABC), Clinical Immunology Laboratory, 821 Príncipe de Gales Avenue, Santo André, São Paulo, Brazil

Full list of author information is available at the end of the article

of fulminant hepatitis is ten times [8–10], particularly when the characteristic Amazon genotypes (HDV-3 and HBV-F) are found combined [11, 12].

An important consideration regarding HDV is the availability of only one treatment option. This happens because this virus does not encode specific enzymes that could be used as target for the development of new antiviral drugs [3]. The efficacy of the treatment done with Interferon-Pegylated, is of only 43 %, and it produces very severe side effects [13].

Another situation that complicates the immune control of this infection is viral recombination. This event can allow the virus to escape the immune pressure imposed by the host's immune system, lead to an increase in viral fitness or even changes in the pathogenicity of the hybrid virus [14]. The recombination can occur when the individual is infected with a mixed viral population, as reported between genotypes 1 and 2 and between genotypes 1 and 4 in Taiwan [15, 16]. The probability of recombination occurring is greater when similar HDV genotypes cause the infection [17–19].

Perhaps because HDV-3 is the most distantly related among all HDVs genotypes, recombination with HDV-3 has never been reported [11, 19]. Furthermore, most studies with HDV-3 analyze only the coding portion of the genome and perhaps analyzing the entire genome, that is a much more detailed approach, may lead to different results in the identification of inter-genotypic recombination.

The adaptation of the virus to the environment can be measured by non synonymous (dN) and synonymous substitutions (dS) present in their genomes. The ratio between dN/dS is used to measure the strength and mode of natural selection acting on protein coding genes [20]. Results  $>1$  indicate that a protein is under positive selection, whereas ratio  $<1$  denotes that the protein is evolving slowly under negative or purifying selection, and  $dN/dS \sim 1$  implies that the protein is evolving neutrally [21].

Another form to better understand an infection is to analyze its diversification, ancestry and phylogeography. These informations can be used to shed light on the demographic, social and biological factors that were responsible for the emergence and spread of the disease present in the population [22, 23]. This approach may also help in the development of preventive strategies against that specific pathogen.

Therefore, according to the above information, this study aims to investigate the presence of recombination and to analyze the epidemiology, ancestry and evolutionary pressures on HDV in samples collected with DBS from the Northern region of Brazil.

## Methods

### Samples and processing

A total of 50 samples from patients in medical care in different regional clinics at the Amazon region of Brazil was collected between 2009 and 2010. The samples were originated from the states of Acre (Cruzeiro do Sul and Rio Branco Cities), Amazonas (Manaus City) and Rondônia (Porto Velho City), all located in the North region of Brazil. The samples were collected and transported in dried-blood spots (DBS 903- Whatman). The mean age of the patients was 32 years (maximum of 69 and minimum of 18), and 52 % were male.

The samples were processed only after the CEP/UNI-FESP (# 72971) was approved, and analyzed anonymously to ensure the protection of the participants. Informed consent was not required because the samples had been previously collected and stored. Patients co-infected with HIV and/or HCV were excluded.

The sample RNA was purified by the method developed by Boom [24], and the cDNA were synthesized using SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA). HDV full genome was sequenced using the shotgun technique, which amplified five overlapping fragments named A - E. The PCR methodology used was based on the study performed by Nakano [25]. The primers used amplified all HDV genotypes, and the sensitivity of amplification was up to  $10^3$  copies/mL. For the D fragment a rescue strategy was necessary, where a nested PCR was implemented, combining the primers 1267\_F with F2 [26] in the 1st round and A3 with 138\_R in the 2nd round [26, 27].

The processing of all samples included a negative control to exclude the possibility of cross contamination and a positive control to ensure the reliability of the reaction. Moreover, both strands of the PCR were sequenced and thereafter a consensus sequence was created, which was used in the analysis to guarantee the best quality of the results obtained.

Sequencing was performed using the *ABI Prism 3130 Genetic Analyzer* (Applied Biosystems, USA). The sequences were edited using the *Sequencher v.4.2 Program* (Gene Code, Ann Arbor, USA) and analyzed using Blastn (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), to verify their similarity with the database sequences.

### Sequence alignment and phylogenetic analyses

The HDV references sequences were obtained from the GenBank database (<http://www.ncbi.nlm.nih.gov/genbank>), described in Table 1. All the sequences were aligned using ClustalX software version 2.1 [28]. After the first alignment, the sequences were manually edited using the SE AL program version 2.0 (<http://tree.bio.ed.ac.uk/software/seal/>).

**Table 1** Hepatitis Delta Virus sequences used as reference

Genotype	Origin	GenBank ID
1a	Iran	AY633627
	Taiwan	AF425644
1c	Somalia	U81988
	Ethiopia	HDU81989
2	Russia	AJ309879
	Japan	AB118846_M37
	Taiwan	X60193
3	Venezuela	AB037947, AB037949, AB037948
	Peru	L22063
4	Japan	AB118842
	Taiwan	AF209859
5	Togo	AM183326
	Senegal	AM183328
	Guinea-Bissau	AM183331
6	Nigeria	AM183329
	Cameroon	AJ584847
7	Cameroon	AM183333, AJ584844
8	Senegal	AM183327
	Congo	AJ584849

Maximum likelihood trees and bootstrap values were obtained using PhyML software [29]. The HKY model and gamma distribution ( $\Gamma$ ) were selected and, based on the likelihood ratio test (LRT), implemented in the jModeltest software [30, 31]. The resulting trees were created and edited using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

### Recombination analyses

The RDP v.4 software was used to determine the presence of recombination in the sequences [32]. 3Seq performs an exact nonparametric test to detect recombination in triplets of sequences [33]. The maximum  $\chi^2$  (MaxChi) is a method implemented by Maynard Smith [34], and it uses variable/invariable sites to detect recombination in pairs of sequences. The maximum match  $\chi^2$  (Chimaera) is a modification of Smith's method that uses variable sites to calculate the maximum  $\chi^2$  match statistics [35]. Geneconv detects recombination by evaluating conserved substitutions in fragments between two sequences [36]. Although evolutionary methods are not explicitly implemented in Geneconv, it is robust and has low levels of false positive [37].

Initially, default parameters were used; then, these were optimized to avoid detection of false positive recombination. In addition, window sizes of 20 to 50 as well as Bonferroni correction were utilized.

### Phylogeographic analysis

A Bayesian Markov chain Monte Carlo (BMCMC) coalescent framework was used to estimate the ancestral genealogy, phylogeographic and time to the most common ancestor (tMRCA). After Markov chain runs enough so that the parameter space have been sufficiently visited, the mean and the 95 % HPD intervals of each parameter was calculated. The HKY model plus a gamma correction ( $\Gamma$ ) was applied to all analyses; the evolutionary and demographic parameters were iteratively adjusted. We used the Bayesian discrete methods to estimate the node probability in regard to the geographical location of samples [38].

In order to explore the correlation between sample location and tree topology, the Bayesian discrete model [38] was used because it allows the reconstruction of ancestral location states while accounting for the uncertainties of maximum posteriori (MP) tree topology and mapping states at nodes.

The Bayesian skyline plot method (BSP) that provides an unbiased description of genetic diversity over time was used (expressed as the product of the effective sample size  $N_e$  and generation time  $\tau$ ). BSP contemplates a wide varied of demographic scenarios without formally assuming any determined model [39]. Since absolute time was used (years) to scale branch lengths and a specific generation time was not assumed, our estimates of  $N_e \tau$  will reflect only the relative genetic diversity of the infections over time.

The Markov Chain Monte Carlo (MCMC) processes were run for  $3 \times 10^8$  generations with the initial 10 % of each run discarded as burn in. The convergence of chains was evaluated using the TRACER software, version 1.5 (available at: <http://beast.bio.ed.ac.uk/>), runs were accepted when all parameters presented the effective sample size number (ESS) greater than 200. Two independent chains were run for each dataset and combined with LogCombiner software [39, 40].

The MP tree was inferred using the BSP method and the maximum clade credibility parameter were annotated using the TreeAnnotator software [39, 40]. All these analyses were performed with the BMCMC approaches implemented in the BEAST package version 1.8.0 [40].

### Estimates of the nonsynonymous (dN) and synonymous (dS) substitution rates

The codon based maximum likelihood method was used to estimate the selective pressures on the PTV sequences. This approach estimates the likelihood of distinct models of codon evolution and computes the ratio ( $dN/dS = \omega$ ) of the number of non-synonymous ( $dN$ ) and synonymous ( $dS$ ) substitution rates between sites considering the phylogenetic relationships of the

sequences [41, 42]. To estimate the overall mean ratio, the one ratio model ( $M_0$ ) was used. This model assumes a single  $\omega$  for all sites in the alignment and is the simplest model. The hypothesis of positive evolution is based on the likelihood ratio test (LRT) comparing pairs of nested models, with the corresponding statistics approximated by a chi square ( $\chi^2$ ) distribution. To detect positive selection, we compared the null hypothesis of model M1, which allows for different proportions of conserved ( $\omega = 0$ ) and neutral codons ( $\omega = 1$ ), against the alternative hypothesis of model M2, which has an additional class of codons with  $\omega$  ratio  $> 1$ , thus assuming positive selection. These models were implemented in the CODEML program of the PAML package, version 4.8, available at <http://abacus.gene.ucl.ac.uk/software/paml.html>.

## Results

### Phylogenetic overview

Out of a total of 50 samples, 46 demonstrated the necessary quality for further analysis. The demographic information and the GenBank ID of each sample used in the analysis are detailed in Table 2 (present in the end of the manuscript). A maximum likelihood tree was initially constructed using the full length genomes of 46 samples and reference genomes of the main genotypes. Figure 1 shows that all sequences generated in the current study belong to genotype 3 and clustered within a single sequence with a bootstrap proportion of 97. Although the samples belong to genotype 3, they formed two groups: one composed of isolates from Yucpa Indians in Venezuela [43] and another composed of Brazilian isolates plus one isolate from the Yavari River in the Amazon basin close to the Peru Brazil border [19]. All sequences used to infer the tree mentioned above lacked a recombination signal.

### Phylogeographic analysis of HDV-3 in the Amazon region

42 samples were used for this analysis, four of the 46 original sequences were excluded because they had significant deletions along the genome. The phylogeography of HDV-3 isolates (Fig. 2), show that Brazilian samples clustered into four clades (named A, B, C, D), one sample from Peru isolated in 1986 is located at the base of these clades. The MP tree presents two large clusters, A and B, containing samples from different geographic locations (Rio Branco, Porto Velho, Cruzeiro do Sul and Manaus) and two small clusters, C and D, with sequences from Rio Branco and Porto Velho, respectively.

The topology also indicates two sets of nearly identical isolates: one from Rio Branco and another from Porto Velho (filled triangles in Fig. 2). The branching pattern of intermingled samples from Rio Branco and Porto

Velho suggests a bidirectional flow of infections between these two locations.

The pattern of geographical dispersal of HDV-3 is depicted by the highest location probability on the nodes of the MP tree (Fig. 2; the location is indicated by colored branches and the location probability by numbers at the nodes). Although the multi-location clades A and B indicate that Rio Branco contributes significantly to the geographical spread of HDV-3, this may occur because samples from this location were overrepresented. When random sets of sequences containing equivalent numbers of samples from Rio Branco and Porto Velho were evaluated, the location probability to support Rio Branco at internal nodes decreased to values below 0.7 (*data not shown*).

### The population dynamics of HDV-3

Figure 3 shows the reconstructed dynamics of the genetic diversity ( $N_e t$ , y axis) plotted against time in years (x axis). The plot shows the mean (thick solid line) and the 95 % highest density interval (thin solid lines). The shape of the plot demonstrates a period of exponential growth that began in the late 1970s and lasted until 1995. The estimated tMRCA values for the HDV-3 and for the Amazon clade were 1934 (1927 until 1940) and 1952 (1948 until 1955), respectively.

### Selective regimen in the large HDV antigen (L-HDAg)

A systematic codon based analysis was performed on HDV isolates using the 215 codons from the L-HDAg region. In Table 3, the estimates of the codon based analysis are summarized. To detect positive selection models M7 and M8 were compared using likelihood ratio test assuming two degrees of freedom between models. Initially an intra-subtype analysis was performed using a dataset with 47 isolates of HDV-3, 46 from this study and one reference sequence from Peru (L22063), results of Model M8 indicated 24 codons in the L-HDAg under positive selection with mean  $\omega$  values of 5.04.

When this same analysis was performed with other HDV genotypes (78 sequences) the Model M8 indicated only nine codons under positive selection, four of which were similar with the result demonstrated with the HDV-3 genotype (8 K, 9 K, 37 T and 120S).

## Discussion

Chronic HDV/HBV infection is perhaps the most intriguing and difficult to treat among the human viral hepatitis and is one of the most neglected diseases worldwide [44]. The study area included was based upon previous articles that demonstrated the endemicity and near exclusivity of HDV infection in the Amazon region of Brazil [12, 19, 45, 46], only two studies have identified

**Table 2** Demographic information and GenBank ID of the samples from our study

Origin	Sample ID	Age	Gender	GenBank ID
Rio Branco (23 samples)	1	36	Male	KF786351
	2 <sup>b</sup>	-	Male	KF786323
	8	43	Male	KF786336
	12	58	Male	KF786343
	13	23	Female	KF786340
	17	48	Female	KF786308
	20	44	Female	KF786341
	21	-	Female	KF786348
	22	41	Male	KF786322
	23	23	Female	KF786334
	25	33	Male	KF786317
	26	39	Female	KF786339
	28	39	Male	KF786309
	30	-	Male	KF786316
	34	58	Male	KF786319
	36	36	Male	KF786324
	41	26	Female	KF786349
	42	23	Female	KF786335
	44	41	Female	KF786342
	45	39	Male	KF786350
	81	24	Female	KF786313
	85	41	Female	KF786330
	88	18	Female	KF786306
Cruzeiro do Sul (2 samples)	3	32	Male	KF786321
	4	42	Male	KF786346
Porto Velho (16 samples)	5	47	Male	KF786327
	6	24	Male	KF786328
	7 <sup>a</sup>	-	Female	KF786331
	9	28	Male	KF786352
	10	31	Female	KF786307
	11	44	Male	KF786332
	14	26	Male	KF786337
	15	33	Male	KF786305
	18	38	Female	KF786325
	19	37	Male	KF786329
	27 <sup>a</sup>	33	Female	KF786311
	29	50	Female	KF786345
	33	69	Female	KF786333
	51 <sup>a</sup>	22	Female	KF786314
	52 <sup>a</sup>	21	Male	KF786318
	60 <sup>a</sup>	61	Male	KF786320

**Table 2** Demographic information and GenBank ID of the samples from our study (Continued)

Manaus (5 samples)	AM1 <sup>b</sup>	-	-	KF786310
	AM2 <sup>b</sup>	-	-	KF786347
	AM3 <sup>b</sup>	-	-	KF786338
	39	57	Male	KF786312
	40	34	Male	KF786326

<sup>a</sup> samples from Porto Velho that are probably indigenous

<sup>b</sup> samples excluded in further analysis

- samples without information

HDV cases outside the locations presented in this study [47, 48].

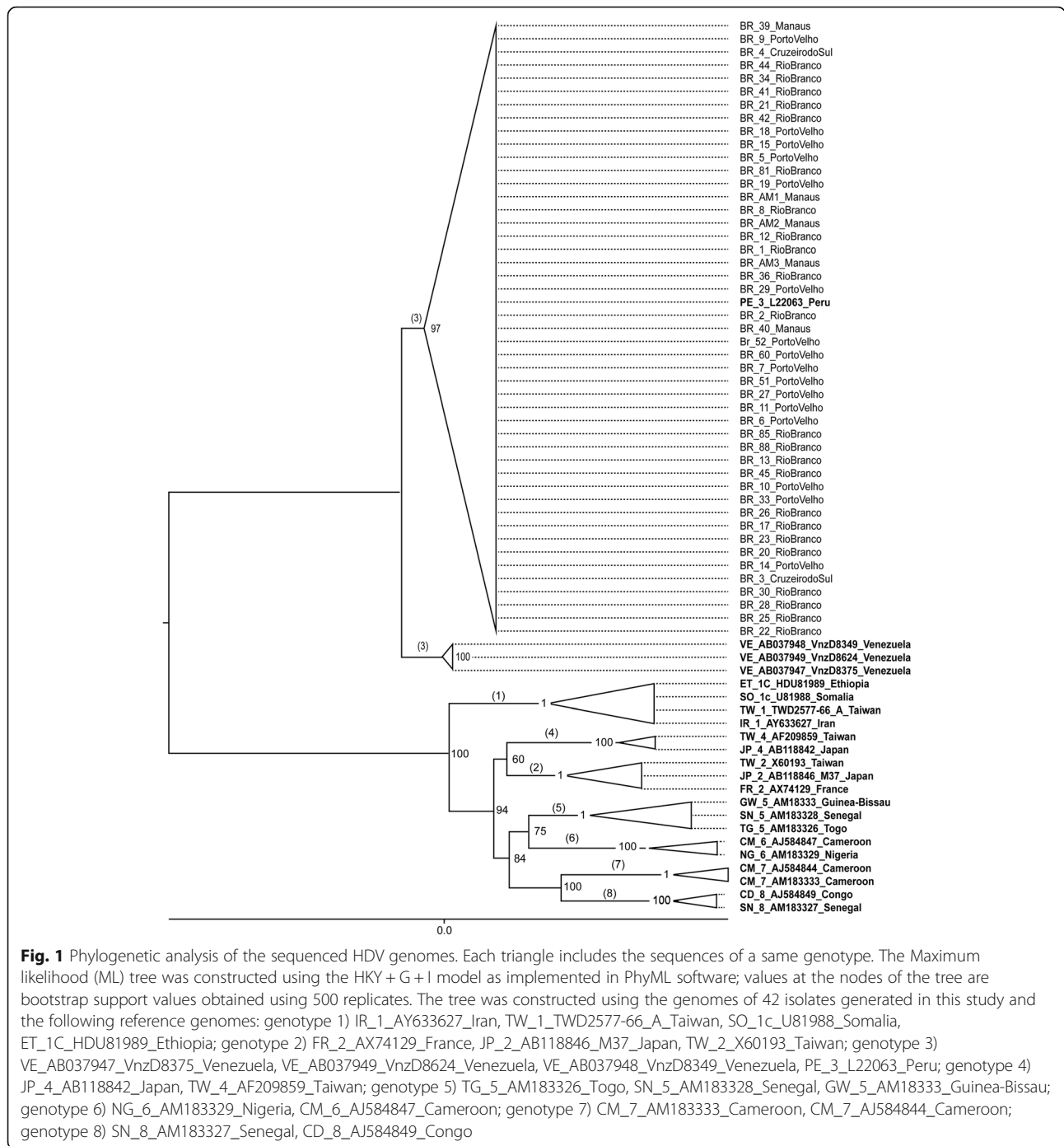
In Brazil, 77 % of the HDV infections occur in the North region [48]. The isolation of this virus in this location can be a consequence that the majority of HDV cases are present in the indigenous population, that generally lives in isolated regions and interacts preferentially among individuals of their own tribes.

The DBS was used to collect the samples because: distant areas were included, this method is less invasive for the patient once the blood is collected via puncture of the finger using a lancet, it requires minimal collector training, allows the samples to be stored at room temperature with transportation in envelopes sent via regular mail (less expensive than using dried ice transportation) [49, 50].

Based upon the phylogenetic tree created from the sequences obtained in this study, it was possible to conclude that all samples belongs to genotype 3, being clustered (97 in the bootstrap) along with one HDV-3 sequence from Peru. These results confirm previous studies, that also demonstrated the high prevalence of this genotype in the studied area [12, 19, 45, 46].

It was not detected any inter-genotypic recombination. This may have happened because we analyzed areas of difficult access, which may have hindered the introduction of different HDV genotypes.

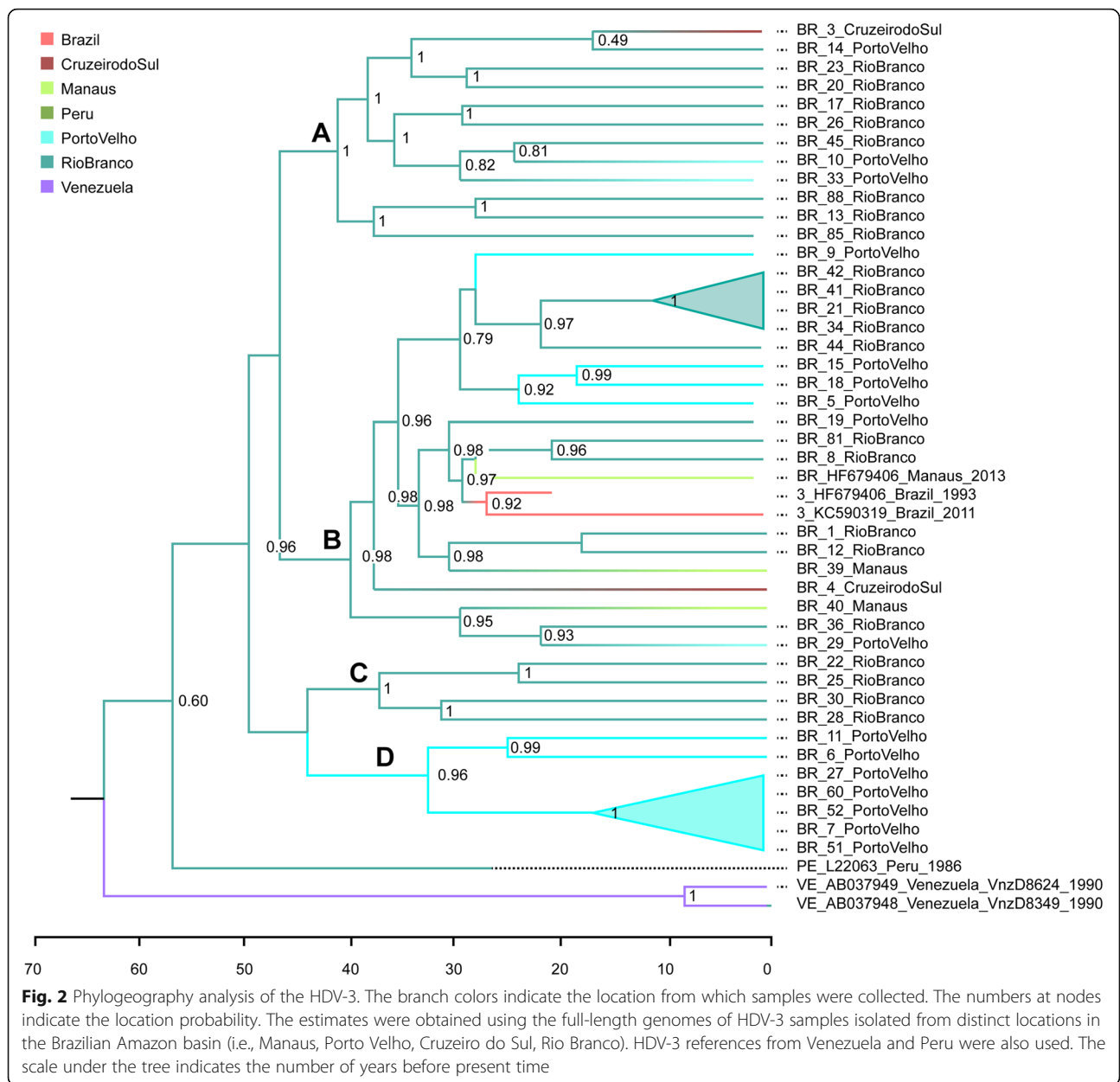
The BMCMC method was used in the phylogeographic analysis because it incorporates the uncertainty of the measurements by considering the intrinsic errors both in the tree reconstruction and in the coalescent method. These considerations guarantee more confiability in the results obtained. This analysis demonstrated the presence of four clades in HDV-3 samples from the Amazon region (A, B, C and D). Two of them (A and B) contained samples from all the locations studied (Rio Branco, Porto Velho, Cruzeiro do Sul and Manaus), whereas clusters C and D were significantly smaller and represented the cities of Rio Branco and Porto Velho, respectively. This analysis also indicates two clusters with nearly identical sequences, one from Rio Branco (composed of four samples) and another from Porto Velho (composed of five samples), which suggests a bidirectional flow of infection in these two cities.



Rio Branco is the capital city of the Acre State, populated in the beginning of the 20th century by natives who already lived in the area, along with the migration of people from Northeast of Brazil and immigrants from Turkey, Portugal and Lebanon attracted by the enrichment of the rubber era [51]. The cluster identified at this location indicates a probable founder effect, that may suggest that the dissemination of this infection was introduced by an indian (HDV-3 is characteristic of the

indigenous population), and not brought with the city foundation.

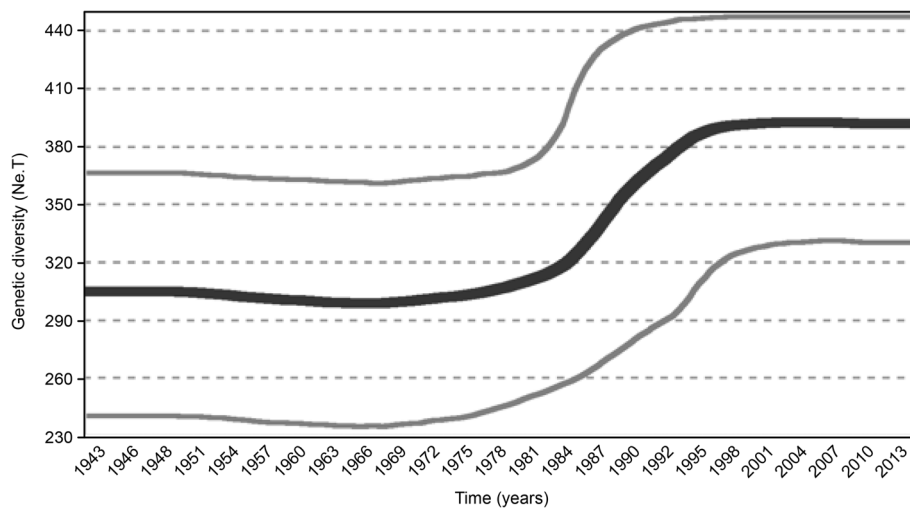
Porto Velho, the capital city of Rondônia State, was founded as a result of the railroad “Madeira Mamoré”, widely used in the rubber and gold eras. Nowadays several indigenous tribes still live in this region [51]. Five samples from this location were probably indigenous (surnames characteristic from tribes living in this area), and they grouped together into a single cluster. This



situation demonstrates that probably a single virus caused the spread of the HDV infection in this region, as well.

These probable founder effects demonstrate the necessity of focusing on prevention strategies against HBV and HDV infection, mainly in the North of Brazil. This region represents respectively 14 and 77 % of the total cases of these two viral infections. The most effective prevention against HDV is the HBV vaccination, since the HDV depends of the HBV for its budding from the infected cell, and if HDV is not contained it can be disseminated to other regions of Brazil, causing a public health problem.

The analysis of the most common ancestor time (tMRCA) showed that the infection of HDV experienced an exponential growth period in the late 1970s that lasted until 1995 and has been maintained since then. In 1989, the HBV vaccination was implemented in the Amazon region and in 1993, the minimal age for vaccination was decreased, these events could explain the plateau in the exponential growth of HDV after 1995; but this same plateau does not exclude the need to focus on preventive campaigns against this infection, since its treatment is not effective, no remission has been observed and the vaccination against HBV is not effective in all vaccinated individuals.



**Fig. 3** Bayesian skyline plot (BSP) of the population dynamics of HDV-3. The plot represents the genetic diversity (y-axis) against chronological time expressed in years (x-axis). The thick line indicates the mean and the thin lines the 95 % confidence posterior interval of the  $N_e t$ . The estimates were obtained using the full-length genomes of 42 HDV-3 isolates

It was also observed that the HDV-3 has more codons under positive selection in the L-HDAg than the other HDV genotypes, but four codons (8 K, 9 K, 37 T and 120S) were present in both analysis indicating that these modifications are not genotype specific. Some modifications were located in probable epitopes of HDV for B lymphocytes (amino acids 3–18; 107–122 and 170–196) [52, 53]. This finding may indicate that the virus is evolving to escape the humoral response of the host.

The variation of epitopes at amino acids 174 to 195, often emerges after severe hepatitis attacks during the chronic phase of infections [15]. Changes in epitopes are common during the chronic phase of infections because of the intense selection forces by the immune defense against the virus [54]. Unfortunately, we did not have information to confirm that these individuals were in this phase of infection.

## Conclusion

Our results confirm the prevalence of HDV-3 in the Amazon region of Brazil. This geographically isolated region may have hindered the introduction of different HDV genotypes, which may account for the absence of inter-genotypic recombination. The analysis of the most common ancestor time showed that the infection with HDV experienced an exponential growth period in the late 1970s that lasted until 1995, after which it has remained constant. Taking into account the probable founder effect established in the cities of Rio Branco and Porto Velho, a focus on preventive methods is recommended against this infection, particularly in the North of Brazil. A positive selection was also found in probable epitopes of HDV for B lymphocytes, which may indicate that the virus is evolving to escape the humoral response of the host.

**Table 3** Codon-based analysis of  $d_N/d_S$  rates in the L-HDAg gene of Hepatitis Delta virus

Model	$\rho$	-lnL	Estimated parameters	Sites <sup>b</sup>
Site-specific models for HDV-3 ( $n = 47$ )				
M7 (beta)	2	5471.30397	$\rho = 0.09695$ $q = 0.15394$ $k$ (ts/tv) = 3.53356	Not allowed
M8 (beta& $\omega$ ) <sup>a</sup>	4	5293.55296	$\rho_0 = 0.82801$ $\rho = 0.11508$ $q = 0.24197$ ( $\rho_1 = 0.17199$ ) $\omega = 5.03937$ $k$ (ts/tv) = 4.64237	4P, 8 K, 9 K, 29E, 37 T, 44 L, 54I, 55 V, 63D, 73P, 77A, 80I, 82S, 84P, 96Q, 120S, 123Q, 130G, 134 K, 143 V, 169 N, 190 T, 198 F, 205H
Site-specific models for the others HDV Genotypes ( $n = 78$ )				
M7 (beta)	2	9893.49320	$\rho = 0.27034$ $q = 0.48352$ $k$ (ts/tv) = 2.60287	Not allowed
M8 (beta& $\omega$ ) <sup>a</sup>	4	9781.56864	$\rho_0 = 0.92957$ $\rho = 0.31395$ $q = 0.63880$ ( $\rho_1 = 0.07043$ ) $\omega = 3.27612$ $k$ (ts/tv) = 2.85894	6 L, 8 K, 9 K, 16I, 37 T, 120S, 157G, 179A, 187I

*lnL* Maximum likelihood of each model

<sup>a</sup> Best-fit model according to the LRT comparing the M7 and M8 models, using two degrees of freedom and a significance level of 0.010

<sup>b</sup> Positive selected codons were identified according to the Bayes Empirical Bayes (BEB) analysis with  $P > 99\%$



### Abbreviations

BMCMC: Bayesian Markov chain Monte Carlo; BSP: Bayesian skyline plot method; cDNA: Complementary DNA; DBS: Dried blood spots; dN: Non synonymous substitutions; dS: Synonymous substitutions; ESS: Effective sample size number; HBV: Hepatitis B Virus; HDV: Hepatitis Delta Virus; HPD: Highest posterior density; L-HDAG: Large HDV antigen region protein; LRT: Likelihood ratio test; MP: Maximum posteriori; PCR: Polymerase chain reaction; tMRCA: Time to the most common ancestor

### Acknowledgements

We would like to thank Dr. Alexandra Reis and Centro de Genomas for their methodological assistance and all the patients who enabled this study.

### Funding

We would like to thank the financial support obtained from São Paulo Research Foundation (FAPESP, grant#2012/13285-0), National Council for Scientific and Technological Development (CNPQ, grant#135382/2011-4).

### Availability of data and material

The datasets generated during and/or analysed during the current study are available in the GenBank repository (<http://www.ncbi.nlm.nih.gov/genbank/>) and the GenBank ID information is present in Table 2.

### Authors' contributions

RSD and SVK were responsible for the conception and designed of the study; MCF and SVK wrote the manuscript; MCF, NMP, LCS, RA, RGA carried out the molecular studies and participated in the sequence alignment; MCF, LCS, ESL participated in the sequence alignment. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

The Ethics approval for this study was number # 72971 (CEP/UNIFESP), and the informed consent was not required because the samples had been previously collected and stored in the laboratory.

### Author details

<sup>1</sup>Retrovirology Laboratory, Federal University of São Paulo, 781 Pedro de Toledo Street, São Paulo, São Paulo, Brazil. <sup>2</sup>Institute of Biological Sciences, Federal University of Pará, 01 Augusto Corrêa Street, Belém, Pará, Brazil. <sup>3</sup>School of Medicine of ABC (FMABC), Clinical Immunology Laboratory, 821 Príncipe de Gales Avenue, Santo André, São Paulo, Brazil.

Received: 23 December 2015 Accepted: 20 September 2016

Published online: 29 September 2016

### References

- Genne D, Rossi I. Hepatitis delta in Switzerland: a silent epidemic. *Swiss Med Wkly*. 2011;141:w13176.
- Cross TJ, Rizzi P, Horner M, Jolly A, Hussain MJ, Smith HM, Vergani D, Harrison PM. The increasing prevalence of hepatitis delta virus (HDV) infection in South London. *J Med Virol*. 2008;80:277–82.
- Wedemeyer H, Manns MP. Epidemiology, pathogenesis and management of hepatitis D: update and challenges ahead. *Nat Rev Gastroenterol Hepatol*. 2010;7:31–40.
- de Truchis P, Le MP, Daou M, Madougou B, Nouhou Y, Moussa Saley S, Sani A, Adehossi E, Rouveix E, Saidou M, et al. High efficacy of first-line ART in a West African cohort, assessed by dried blood spot virological and pharmacological measurements. *J Antimicrob Chemother*. 2016.
- Gueye SB, Diop-Ndiaye H, Diallo MM, Ly O, Sow-Ndoye A, Diagne-Gueye ND, Kebe-Fall K, Diop F, Gaye-Diallo A, Belec L, et al. Performance of Roche CAP/CTM HIV-1 qualitative test version 2.0 using dried blood spots for early infant diagnosis. *J Virol Methods*. 2016;229:12–5.
- Wei X, Smith AJ, Forrest DW, Cardenas GA, Beck DW, LaLota M, Metsch LR, Sionean C, Owen SM, Johnson JA. Incident infection and resistance mutation analysis of dried blood spots collected in a field study of HIV risk groups, 2007–2010. *PLoS One*. 2016;11:e0159266.
- Rizzetto M, Canese MG, Arico S, Crivelli O, Trepo C, Bonino F, Verme G. Immunofluorescence detection of new antigen-antibody system (delta/anti-delta) associated to hepatitis B virus in liver and in serum of HBsAg carriers. *Gut*. 1977;18:997–1003.
- Smedile A, Farci P, Verme G, Caredda F, Cargnel A, Caporaso N, Dentico P, Trepo C, Opolon P, Gimson A, et al. Influence of delta infection on severity of hepatitis B. *Lancet*. 1982;2:945–7.
- Lai MM. The molecular biology of hepatitis delta virus. *Annu Rev Biochem*. 1995;64:259–86.
- Purcell RH, Gerin JL. Hepatitis Delta virus. In: Fields BN, Knipe DM, Howley PM, editors. *Fields virology*. 3rd ed. Philadelphia: Lippincott-Raven; 1996. p. 2819–29.
- Casey JL, Niro GA, Engle RE, Vega A, Gomez H, McCarthy M, Watts DM, Hyams KC, Gerin JL. Hepatitis B virus (HBV)/hepatitis D virus (HDV) coinfection in outbreaks of acute hepatitis in the Peruvian Amazon basin: the roles of HDV genotype III and HBV genotype F. *The Journal of infectious diseases*. 1996;174:920–6.
- Quintero A, Uzcategui N, Loureiro CL, Villegas L, Illarramendi X, Guevara ME, Ludert JE, Blitz L, Liprandi F, Pujol FH. Hepatitis delta virus genotypes I and III circulate associated with hepatitis B virus genotype F in Venezuela. *J Med Virol*. 2001;64:356–9.
- Han Z, Nogusa S, Nicolas E, Balachandran S, Taylor J. Interferon impedes an early step of hepatitis delta virus infection. *PLoS One*. 2011;6:e22415.
- Chao M. RNA recombination in hepatitis delta virus: implications regarding the abilities of mammalian RNA polymerases. *Virus Res*. 2007;127:208–15.
- Wu JC, Chiang TY, Shiue WK, Wang SY, Sheen IJ, Huang YH, Syu WJ. Recombination of hepatitis D virus RNA sequences and its implications. *Mol Biol Evol*. 1999;16:1622–32.
- Wang TC, Chao M. RNA recombination of hepatitis delta virus in natural mixed-genotype infection and transfected cultured cells. *J Virol*. 2005;79:2221–9.
- McCahon D, Slade WR, Priston RA, Lake JR. An extended genetic recombination map for foot-and-mouth disease virus. *J Gen Virol*. 1977;35:555–65.
- McCahon D, King AM, Roe DS, Slade WR, Newman JW, Cleary AM. Isolation and biochemical characterization of intertypic recombinants of foot-and-mouth disease virus. *Virus Res*. 1985;3:87–100.
- Casey JL, Brown TL, Colan EJ, Wignall FS, Gerin JL. A genotype of hepatitis D virus that occurs in northern South America. *Proc Natl Acad Sci U S A*. 1993;90:9016–20.
- Jeffares DC, Tomiczek B, Sojo V, Dos Reis M. A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. *Methods Mol Biol*. 2015;1201:65–90.
- Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*. 1977;267:275–6.
- Markov PV, Pepin J, Frost E, Deslandes S, Labbe AC, Pybus OG. Phylogeography and molecular epidemiology of hepatitis C virus genotype 2 in Africa. *The Journal of general virology*. 2009;90:2086–96.
- Lago BV, Mello FC, Kramvis A, Niel C, Gomes SA. Hepatitis B virus subgenotype A1: evolutionary relationships between Brazilian, African and Asian isolates. *PLoS One*. 2014;9:e105317.
- Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, van der Noordaa J. Rapid and simple method for purification of nucleic acids. *J Clin Microbiol*. 1990;28:495–503.
- Nakano T, Lu L, Hu X, Mizokami M, Orito E, Shapiro C, Hadler S, Robertson B. Characterization of hepatitis B virus genotypes among Yucpa Indians in Venezuela. *The Journal of general virology*. 2001;82:359–65.
- Behzadian F, Sabahi F, Karimi M, Sadeghizadeh M, Maghsoudi N, Forooshani RS, Shahinsaz L. Molecular phylogenetic analysis of Iranian HDV complete genome. *Virus Genes*. 2005;30:383–93.
- Ma SP, Sakugawa H, Makino Y, Tadano M, Kinjo F, Saito A. The complete genomic sequence of hepatitis delta virus genotype 1b prevalent in Okinawa, Japan. *The Journal of general virology*. 2003;84:461–4.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947–8.
- Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003;52:696–704.
- Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics*. 1998;14:817–8.

31. Posada D. jModelTest: phylogenetic model averaging. *Mol Biol Evol.* 2008;25:1253–6.
32. Martin DP, Williamson C, Posada D. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics.* 2005;21:260–2.
33. Boni MF, Posada D, Feldman MW. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics.* 2007;176:1035–47.
34. Smith JM. Analyzing the mosaic structure of genes. *J Mol Evol.* 1992;34:126–9.
35. Posada D, Crandall KA. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci U S A.* 2001;98:13757–62.
36. Sawyer S. Statistical tests for detecting gene conversion. *Mol Biol Evol.* 1989;6:526–38.
37. Bay RA, Bielawski JP: Recombination Detection Under Evolutionary Scenarios Relevant to Functional Divergence. *J Mol Evol.* 2012.
38. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. *PLoS Comput Biol.* 2009;5:e1000520.
39. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol.* 2005;22:1185–92.
40. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012;29:1969–73.
41. Yang Z, Nielsen R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol.* 2002;19:908–17.
42. Yang Z, Nielsen R, Goldman N, Pedersen AM. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics.* 2000;155:431–49.
43. Nakano T, Shapiro CN, Hadler SC, Casey JL, Mizokami M, Orito E, Robertson BH. Characterization of hepatitis D virus genotype III among Yucpa Indians in Venezuela. *The Journal of general virology.* 2001;82:2183–9.
44. Braga WS, de Oliveira CM, de Araujo JR, Castilho Mda C, Rocha JM, Gimaque JB, Silva ML, Vasconcelos HL, Ramasawmy R, Parana R. Chronic HDV/HBV co-infection: predictors of disease stage—a case series of HDV-3 patients. *J Hepatol.* 2014;61:1205–11.
45. Gomes-Gouvea MS, Soares MC, Bensabath G, de Carvalho-Mello IM, Brito EM, Souza OS, Queiroz AT, Carrilho FJ, Pinho JR. Hepatitis B virus and hepatitis delta virus genotypes in outbreaks of fulminant hepatitis (Labrea black fever) in the western Brazilian Amazon region. *The Journal of general virology.* 2009;90:2638–43.
46. Gomes-Gouvea MS, Pereira Soares Mdo C, de Carvalho Mello IM G, Brito EM, Pereira Moia Lde J, Bensabath G, Nunes HM, Carrilho FJ, Pinho JR. Hepatitis D and B virus genotypes in chronically infected patients from the Eastern Amazon Basin. *Acta Trop.* 2008;106:149–55.
47. Barros LM, Gomes-Gouvea MS, Pinho JR, Alvarado-Mora MV, Dos Santos A, Mendes-Correa MC, Caldas AJ, Sousa MT, Santos MD, Ferreira AS. Hepatitis Delta virus genotype 8 infection in Northeast Brazil: inheritance from African slaves? *Virus Res.* 2011;160:333–9.
48. Mendes-Correa MC, Gomes-Gouvea MS, Alvarado-Mora MV, Da Silva MH, Lazari C, Cavalcanti NC, Alonso FK, Carpinelli CC, Uip DE, Pinho JR. Hepatitis delta in HIV/HBV co-infected patients in Brazil: is it important? *Int J Infect Dis.* 2011;15:e828–832.
49. Garcia Boy R, Henseler J, Mattern R, Skopp G. Determination of morphine and 6-acetylmorphine in blood with use of dried blood spots. *Ther Drug Monit.* 2008;30:733–9.
50. la Marca G, Malvagia S, Filippi L, Luceri F, Moneti G, Guerrini R. A new rapid micromethod for the assay of phenobarbital from dried blood spots by LC-tandem mass spectrometry. *Epilepsia.* 2009;50:2658–62.
51. Instituto Brasileiro de Geografia e Estatística (IBGE): Estimativa populacional 2011. Prefeitura Municipal de Porto Velho: Instituto Brasileiro de Geografia e Estatística (IBGE); 2011.
52. Nisini R, Paroli M, Accapezzato D, Bonino F, Rosina F, Santantonio T, Sallusto F, Amoroso A, Houghton M, Barnaba V. Human CD4+ T-cell response to hepatitis delta virus: identification of multiple epitopes and characterization of T-helper cytokine profiles. *J Virol.* 1997;71:2241–51.
53. Wang JG, Jansen RW, Brown EA, Lemon SM. Immunogenic domains of hepatitis delta virus antigen: peptide mapping of epitopes recognized by human and woodchuck antibodies. *J Virol.* 1990;64:1108–16.
54. Wang SY, Wu JC, Chiang TY, Huang YH, Su CW, Sheen IJ. Positive selection of hepatitis delta antigen in chronic hepatitis D patients. *J Virol.* 2007;81:4438–44.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

