Article

Molecular Epidemiology of HCV Infection among Multi-Transfused β-Thalassemia Patients in Eastern India: A Six-Year Observation

Supradip Dutta [1,†] , Aritra Biswas [1,†] , Sagnik Bakshi [1] , Promisree Choudhury [1] , Raina Das [1] , Shreyasi Nath [1] , Prosanto Chowdhury [2] , Maitreyee Bhattacharyya [3] , Sharmistha Chakraborty [4] , Shanta Dutta [1] and Provash Chandra Sadhukhan [1,*]

1   Indian Council of Medical Research-National Institute of Cholera and Enteric Diseases P-33, Scheme XM, CIT Road, Beliaghata, Kolkata 700010, India; supradipdutta.sd@gmail.com (S.D.); aritrabiswas7@gmail.com (A.B.); bakshi445@gmail.com (S.B.); promisree@gmail.com (P.C.); raina2201@gmail.com (R.D.); shreyasiicmr29@gmail.com (S.N.); drshantadutta@gmail.com (S.D.)
2   Peerless Hospital & B. K. Roy Research Centre 360 Panchasayar, Kolkata 700094, India; drprosanto.chowdhury@gmail.com
3   Department of Haematology, Kolkata Medical College & Hospital, 88 College Street, Kolkata 700012, India; mbhattyacharyaa@yahoo.co.in
4   Thalassaemia Unit with Pathology Department, Burdwan Medical College and Hospital Baburbag, Rajbati, Burdwan 713104, India; thal.burdwan@gmail.com
*   Correspondence: provash2000@gmail.com or sadhukhan.pc@icmr.gov.in; Tel.: +91-98305-46338
†   These authors contributed equally to this work.

Abstract: Background: HCV infection is very common in multi-transfused β-thalassemia patients who need regular blood transfusions. **Aim:** The study was conducted to determine the epidemiology of HCV in multi-transfused β-thalassemia patients in West Bengal, India. **Methods:** Over a span of six years, blood samples were collected from HCV sero-reactive β-thalassemia patients and processed for viral RNA isolation followed by nested RT-PCR for qualitative viremia detection. The HCV genotype was determined by amplifying the partial HCV core gene by nested RT-PCR followed by DNA sequencing and NCBI genotyping tools. Phylogenetic and phylogeographic studies were performed with MEGA-X and BEAST software, respectively. **Results:** Out of 917 multi-transfused HCV sero-reactive β-thalassemia patients, 598 (65.21%) were HCV RNA positive while 250 (41.80%) had spontaneously cleared the virus. A significant percentage of male patients from rural areas (*p* = 0.042) and economically backward class (*p* = 0.002) were at higher risk of HCV infection. Female thalassemia patients and individuals belonging to ages 11–15 years had higher chances of spontaneous clearance. The most prevalent circulatory HCV genotype was 3a (78.26%) followed by 1b (12.04%). Phylogeographic analyses revealed that the 3a strains share genomic similarities with strains from Pakistan, Sri Lanka, and Thailand, whereas the 1b strains share similarities with strains from Thailand, Vietnam, Russia, and China. Uncommon HCV subtypes 3g and 3i were also detected. **Conclusion:** The high prevalence of HCV infection among β-thalassemia patients of West Bengal, India indicates NAT-based assays should be implemented for HCV screening in donor blood to eliminate HCV by 2030.

Keywords: hepatitis C virus; molecular epidemiology; thalassemia; West Bengal (India); nosocomial infection

1. Introduction

HCV is an enveloped, positive sense, ~9.6 kb ssRNA virus belonging to the Hepaciviral genus and *Flaviviridae* family that can cause chronic liver diseases, including cirrhosis of the liver and hepatocellular carcinoma (HCC). HCV mainly transmits through blood, blood products, and body fluids infecting an estimated 200 million people globally [1].After

entering the blood stream, the virus makes its way into the liver and replicates in hepato-cytes, resulting in acute or chronic liver infection. Based on viral genetic diversity, HCV has been classified into seven genotypes (1–7) and over sixty-seven subtypes [2]. Before 2011, Ribavirin (RIB) plus Pegylated interferon (Peg-IFN) was the gold standard for HCV treatment. However, a lower sustainable viral response (SVR) and adverse side effects to patients lead to the introduction of direct acting antivirals (DAAs) [3]. According to the treatment guidelines by the American Association for the Study of Liver Diseases (AASLD) and also the Indian National guidelines for the management of viral hepatitis, determination of the HCV genotype is important to manage DAAs failure cases [4–6]. HCV genotypes differ from each other not only in nucleotide sequences but also in geographical distribution [7].

β-thalassemia is an inherited autosomal recessive disorder that arises from reduced or lack of synthesis of the beta-globin chain of hemoglobin [8]. β-thalassemia is prevalent all over the world; however, the frequency is higher in the Mediterranean region, the Middle East, the Indian subcontinent, and Southeast Asian countries [9]. The burden of hemoglobinopathies, an inherited genetic disorder, is frequent in India because a vast majority of ethnic groups still follow endogamy. Due to the diversified population, the prevalence of β-thalassemia and other hemoglobinopathies varies with different castes/ethnic groups [10]. Approximately one-tenth of the world's thalassemic patients are born in India every year of which about 1–3% in southern India and 3–15% in northern India are carriers [11]. A lack of carrier screening and prenatal diagnosis is one of the main reasons why India has such a high number of thalassemia-positive individuals, especially children who go undiagnosed for a long time. Many parents do not even know that they are thalassemia carriers till they unintentionally pass on the gene to their children [12]. Regular blood transfusion is the major treatment option for multi-transfused β-thalassemia patients; these patients are at a higher risk of acquiring post-transfusion HCV infection, if the transfused blood was collected during the donor's seronegative "window" period [13]. HCV infection in β-thalassemia patients may lead to HCC which can further deteriorate the situation [14]. Moreover, infection management in β-thalassemia patient is difficult because they are immunocompromised [15].

The genomic diversity of HCV depends on geographical localization. Different countries may have different genotype distributions [16]. In India, genotype 3 is prevalent in the north, east, and west, and genotype 1 in the south [17]. Although a hospital-based study on the prevalence and genotypic distribution of HCV in thalassemia patients has been conducted [18], there was still a lack of a more comprehensive study on the genomic diversity of HCV in β-thalassemic individuals from India. This study was undertaken to highlight the viremia, genomic diversity, and evolution dynamics of HCV among the multi-transfused β-thalassemia patient population in the Indian state of West Bengal from 2014 to 2019.

2. Materials and Methods

2.1. Ethical Statement

Written informed consent was obtained from the patients before including them in this study. This study protocol complied with the Helsinki Declaration of 2013, and was approved by the Institutional Ethical Committee, the Indian Council of Medical Research-National Institute of Cholera and Enteric Diseases (ICMR-NICED), Kolkata, India.

2.2. Study Design

A total of 917 HCV sero-reactive β -thalassemic patients were enrolled in this prospective study (from January 2014 to December 2019). Patients' blood samples were collected by venipuncture in clot vials followed by the collection of demographic and clinical data from 10 collaborating transfusion centers (TCs). Patients with other viral co-infections were not included in this study. These 10 TCs were named TC 1–10 and are located around ten southern districts namely, Purulia, Bankura, Burdwan, East Midnapur, West Midnapur,

Howrah, Hooghly, North and South 24 parganas and Kolkata in West Bengal, the eastern state of India (Figure 1a). All laboratory work was performed at ICMR-NICED, Kolkata. HCV seroreactivity was rechecked by HCV ELISA (HCV Ag-Ab Monolisa kit; BioRad, Mames-la-coquette, France). HCV RNA-positive patients were classified into seven age groups (AG), 1 to 7, depending on their age, namely, AG-1 (1–5 years), AG-2 (6–10 years), AG-3 (11–15 years), AG-4 (16–20 years), AG-5 (21–25 years), AG-6 (26–30 years), and AG-7 (30 years and above). Patients' ages ranged from 2 to 35 years with a mean (±SD) of 11.5 ± 7.5 years. Spontaneous clearance was defined as the undetectable level of HCV RNA (viral load < 50 IU/mL) by quantitative real-time PCR in two consecutive testing 6 months apart. Patients who achieved spontaneous clearance in this study were followed up for 2 years after the initial detection of infection.
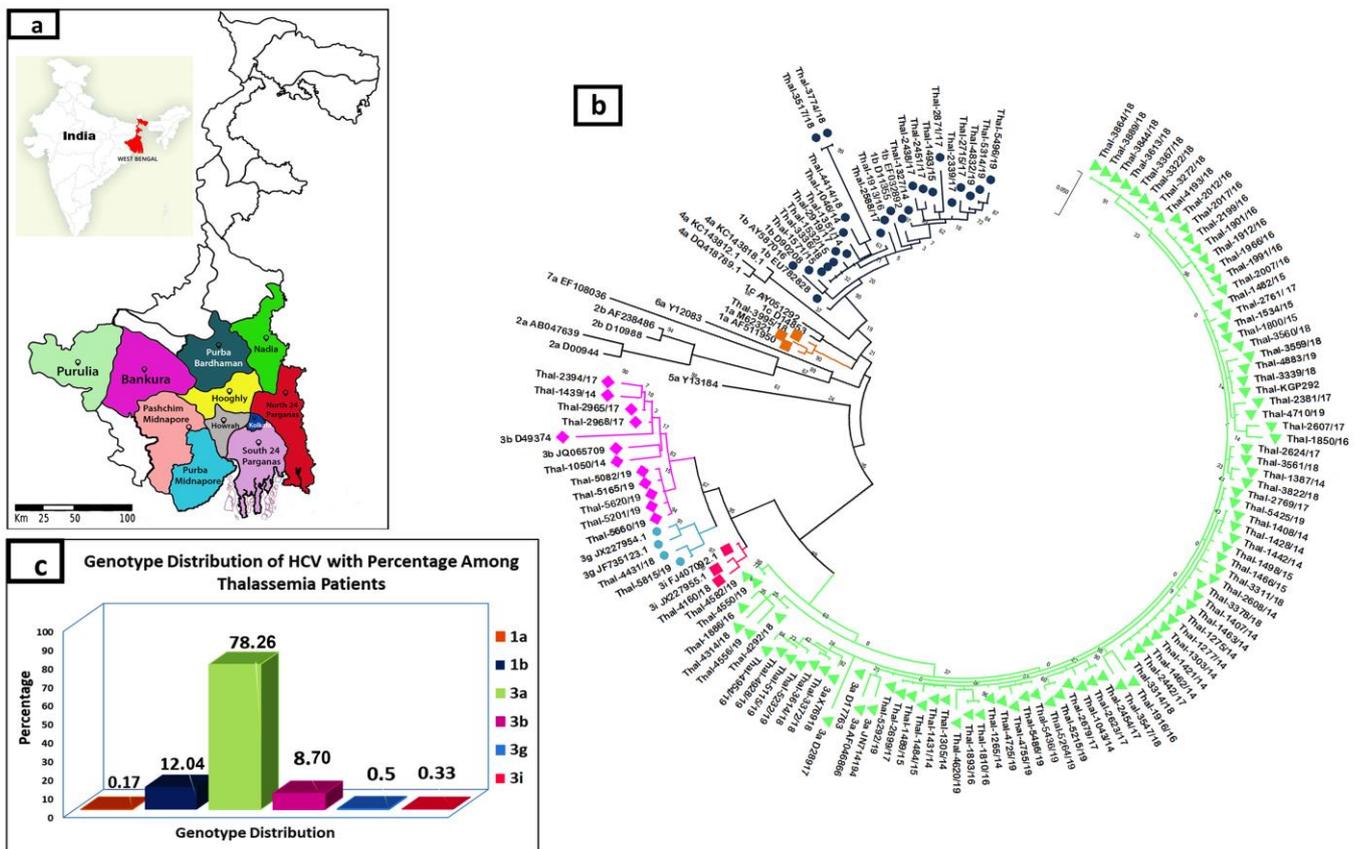
Figure 1. (**a**): Map of relative positions of transfusion centers in West Bengal (Map courtesy: d-maps.com). (**b**): Phylogenetic tree was constructed by MEGA X software package with 122 representative sequences and 30 reference sequences downloaded from NCBI. Reference sequences used were 1a (M62321, AF511950); 1b (EF032892, D11355, D90208, EU782828, AY567016); 1c (AY061292, D14853); 2a (AB047639, D00944); 2b (AF238486, D10988); 3a (X76918, D28917, D17763, AF046866, JN714194); 3b (D49374, JQ065709), 3g (JX227954, JF735123); 3i (FJ407092, JX227955); 4a (KC143818, KC143812, DQ418789); 5a (Y13184); 6a (Y12083); 7a (EF108036); (**c**): Graphical representation HCV genotype distribution in thalassemia patients, based on a percentage.

2.3. Detection of HCV RNA and Genotyping

Viral RNA was isolated from all HCV sero-reactive serum samples using the QIAamp viral RNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. In brief, viral RNA was extracted from 140 µL serum, eluted in 50 µL elution buffer, and stored at −80 °C for further use.

Detection of HCV viral RNA was performed by nested RT-PCR based on 5′UTR of the HCV genome as described previously [19]. In brief, the first round one tube RT-PCR

was completed in 20 µL total reaction volume containing 2 µL isolated RNA and the second round nested PCR was performed in 25 µL total volume using 2 µL of 1st round RT-PCR product. A positive band at 256 bp in 1.5% agarose gel stained with ethidium bromide was observed using a gel documentation system (BioRad, Hercules, CA, USA). Quantitative HCV RNA was estimated using a Qiagen quantitative real-time RT-PCR (qRT-PCR) kit (QuantiFast Pathogen RT-PCR plus IC Kit, Germany). The HCV primer and probe sequences were directed against the 5′UTR of the HCV genome [20] using the 4th WHO International Standard for HCV (NIBSC code 06/102) as standard.

For the determination of the HCV genotype, nested RT-PCR amplified amplicons of the partial HCV core gene (405 bp) [19] were gel purified and directly used for DNA sequencing analyses in an automated DNA sequencer, model 3130XL (ABI, Van Allen way, Carlsbad, CA, USA) using a Big Dye terminator 3.1 kit (Applied Biosystems, Austin, Texas, USA). HCV genotyping was determined by NCBI genotyping tool, available at "https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi" (accessed on 20 February 2020).

2.4. Phylogenetic Analysis

Phylogenetic analysis was performed using 122 representative partial core sequences from HCV RNA positive samples and 29 reference sequences from NCBI. To investigate the evolutionary linkage among lab strains and reference strains, partial core sequences of eighty-six laboratory-isolated 3a strains (Accession Numbers, MN635186-MN635242, MW330397-330425), twenty-six 1b strains (Accession Numbers, MN617343-MN617360, MW368893-MW368895), ten 3b strains (Accession Numbers, MN605946-MN605950, MW368887-MW368891), two 3g strains (Accession numbers—MW368883, MW368885), one 3i strain (Accession number—MW368886), and one 1a strain (Accession number-MW368892) were aligned with HCV reference strains using the Molecular Evolutionary Genetics Analysis tool (MEGA-X) [21]. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura–Nei models [22]. This analysis involved 151 nucleotide sequences. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). There was a total of 300 positions in the final dataset.

2.5. Phylogeographic Analysis

Bayesian Evolutionary Analysis Sampling Trees (BEAST), package 1.10.4 available at "http://beast.community/programs" (10 February 2020) was used to study the phylogeographic relationship and distribution of HCV between neighboring countries. HCV core sequences from the following countries, India, China, Pakistan, Bangladesh, Myanmar, Thailand, Japan, Vietnam, Iran, and Australia were downloaded from the ViPR (Virus Pathogen Resource) database, available at "www.viprbrc.org/brc/home.spg?decorator=vipr" (20 February 2020); the HCV sequence database available at "https://hcv.lanl.gov/content/index" (21 February 2020), and NCBI. The first 300 bp (Respective of 371 nt to 671 nt of the strain H77-accession number NC-004102) of conserved HCV core regions were used for analysis. The General Time Reversible (GTR) model of nucleotide substitution was used with gamma + invariant sites distribution. For the clock setting, an uncorrelated relaxed clock type was used [23] along with the Bayesian Skyline Coalescent model. It also ensured the achievement of effective sampling sizes (ESS) > 200 in TRACER, available at "http://beast.community/tracer" (10 February 2020) [24]. The results of BEAST analyses were evaluated in TRACER. Tree Annotator, Available at "https://beast.community/treeannotator" (10 February 2020) and Fig-Tree, available at "https://beast.community/figtree" (10 February 2020) software were used to construct the final tree.

2.6. Statistical Analysis

Paired *t*-tests were performed to interpret if there was any significant association between male and female viremia among different age groups, demographic distributions, and socio-economic parameters. For genotype distribution, the multivariate one-way ANOVA was used to see if there was any relation between male and female patients with HCV genotypes across the age groups or other factors. The Cochran–Mantel–Haenszel (CMH) test was used to evaluate spontaneous clearance between males and females with respect to age groups. Statistical significance was defined as a *p*-value of less than 0.05.

3. Results

3.1. Demographic Data and Distribution of HCV Viremia

A total of 917 HCV sero-reactive serum samples (561 male and 356 female) were collected over six years from multi-transfused β-globulin synthesis defective thalassemia patients from 10 TCs located around ten districts of West Bengal, India. Out of 917 HCV sero-reactive samples, 598 (65.21%) were found to be HCV RNA positive. Out of these 598 HCV RNA positive individuals, 63.71% (*n* = 381) were male and 36.29% (*n* = 217) were female, respectively. Overall, HCV RNA positivity among males was significantly higher than the females (Table 1). It was also observed that HCV prevalence was significantly higher in rural areas (*p* = 0.042) and among poor populations (*p* = 0.002) (Table 1).

Table 1. Distribution of HCV viremia among various demographic and socio-economic groups of β-thalassemia patients.

HCV Positive (*N* = 598)		HCV Viremia (65.21%)		
Variables		(*n* = 381) Male (%)	(*n* = 217) Female (%)	*p*-Value
Age group	1 to 5	67.69	60.24	
	6 to 10	75.78	69.31	
	11 to 15	62.78	54.62	
	16 to 20	62.26	60	0.008 *
	21 to 25	65.22	64.29	
	26 to 30	71.43	57.14	
	>30	71.43	57.14	
Transfusion centers	TC-1	63.33	55	
	TC-2	76	73.33	
	TC-3	68	58.82	
	TC-4	67.16	60.93	
	TC-5	67.21	62.16	<0.0001 *
	TC-6	71.42	64.28	
	TC-7	64.86	57.44	
	TC-8	73.91	73.33	
	TC-9	67.39	58.06	
	TC-10	71.15	60	
Locality of Residence	Rural	69.9	63.3	0.042 *
	Urban	60.5	52.56	
Economic class	Above poverty level	58.18	51.38	0.002 *
	Below poverty level	70.22	63.5	

* Statistically significant.

3.2. Distribution of HCV Genotype

A total of 598 partial core gene sequences (405 bp) were amplified and sequenced for HCV genotyping and phylogenetic analyses. Sequences were aligned with reference sequences by the NCBI genotyping tool and it revealed that 87.80% ($n = 525$) patients of our study population were infected with HCV genotype 3 and 12.20% ($n = 73$) patients were infected with HCV genotype 1 (Table 2). The circulating HCV subtypes that were detected in this region were: 1a, 1b, 3a, 3b, 3g, and 3i with subtype 3a being the predominant subtype (78.26%, $n = 468$) followed by subtype 1b (12.04%, $n = 72$) and 3b (8.70%, $n = 52$) (Table 2), while subtypes 3g, 3i and 1a were found in only three (0.50%), 2 (0.33%) and one (0.17%) patient(s) respectively. Studies on the distribution of HCV genotypes in each TC individually revealed that HCV genotype 3a was detected and found significantly high in all the 10 TCs (TC1-TC10) (Table 2). HCV genotype 3a was also predominant in rural areas (80%) and also among poor people (80.66%). Population-based DNA sequencing data from the partial core region of the HCV genome showed no multiple HCV genotype infections in any of the β-thalassemia patients in this study population. HCV sequences reported in this study were deposited in GenBank with accession numbers as mentioned in the Materials and Methods section.

Table 2. Gender-wise HCV genotype distribution, demographic data, and socio-economic status of research participants with β-Thalassemia.

HCV Positive [+] = 592		HCV Genotypes						
	Variables	Genotype 1b ($N = 72$)		Genotype 3a ($N = 468$)		Genotype 3b ($N = 52$)		*p*-Value
		($n = 50$) Male (%)	($n = 22$) Female (%)	($n = 295$) Male (%)	($n = 173$) Female (%)	($n = 33$) Male (%)	($n = 19$) Female (%)	
Age-group	1 to 5	2.17	1.45	54.35	31.88	5.80	4.35	0.0007 *
	6 to 10	2.11	1.05	55.26	33.16	5.26	3.16	
	11 to 15	13.07	5.11	46.59	27.28	5.11	2.84	
	16 to 20	20.84	8.33	39.58	22.92	6.25	2.08	
	21 to 25	31.82	13.64	31.82	18.18	4.54	0.00	
	26 to 30	11.11	11.11	44.45	22.22	11.11	0.00	
	>30	22.22	11.11	33.34	11.11	11.11	11.11	
Transfusion Centre	TC 1	10	3.33	50	26.67	6.67	3.33	<0.0001 *
	TC 2	6.67	3.33	50	30	6.67	3.33	
	TC 3	11.11	3.70	48.15	25.93	7.41	3.70	
	TC 4	8.53	4.27	49.29	29.38	4.74	3.79	
	TC 5	9.38	3.13	50	28.13	6.25	3.13	
	TC 6	8.33	4.17	50	29.17	4.17	4.17	
	TC 7	4.05	1.35	54.05	32.43	5.41	2.70	
	TC 8	11.11	3.70	48.15	25.93	7.41	3.70	
	TC 9	8.51	4.26	48.94	29.79	6.38	2.13	
	TC 10	10.34	5.17	48.28	29.31	5.17	1.72	
Locality of residence	Rural	7.63	3.30	50.52	29.48	5.57	3.50	0.0519
	Urban	12.15	5.61	46.72	28.03	5.60	1.89	
Economic class	Above poverty level	7.13	3.05	50.91	29.74	5.91	3.26	0.0898
	Below poverty Level	14.86	6.93	44.55	26.73	3.96	2.97	

* Statistically significant; [+] Since genotypes 3g, 3i and 1a were very low in numbers they were not included in this statistical analysis table to remove unnecessary bias.

3.3. HCV Spontaneous Clearance

Out of 598 HCV RNA-positive patients, 485 patients (309 males and 176 females) could be followed up for two years from the day they tested positive. It was observed that female patients could spontaneously clear the HCV virus more efficiently than male patients (Table 3) and age group 11–15 years could clear the HCV infection more efficiently than other age groups (Table 3). As per the CMH test, female patients in the age group 11–15 years could more efficiently clear the virus than other groups (Table 3).

Table 3. Spontaneous HCV clearance in β-thalassemia patients of various age and sex groups.

Gender	Male (Total Spontaneous Clearance: Followed up: Total = 136:309:381)		Female Total Spontaneous Clearance: Followed up: Total = 114:176:217)		*p*-Value
Age Group	Spontaneous Clearance	No Clearance	Spontaneous Clearance	No Clearance	
1 to 5	30	42	25	18	
6 to 10	45	55	39	19	
11 to 15	52	47	43	10	<0.001 *
16 to 20	6	19	5	10	
21 to 25	3	10	2	5	

* Statistically significant.

3.4. Phylogeography of HCV Isolates

To examine the phylogeographic similarity between India and its neighboring countries, a retrospective Bayesian-dated phylogeographic analysis was performed using partial core gene sequences of HCV strains obtained from NCBI sources available online. The maximum clade credibility tree was constructed on the basis of partial core gene sequences for genotypes 3a and 1b. We found that the majority of HCV genotype 3a sequences isolated in our laboratory formed three distinct clusters resulting in the Pakistan clade, the Thailand clade, and the India/Sri Lanka clade (Figure 2) whereas laboratory-isolated HCV genotype 1b strains were similar to isolates from Thailand, Vietnam, Russia, and China (Figure 3).
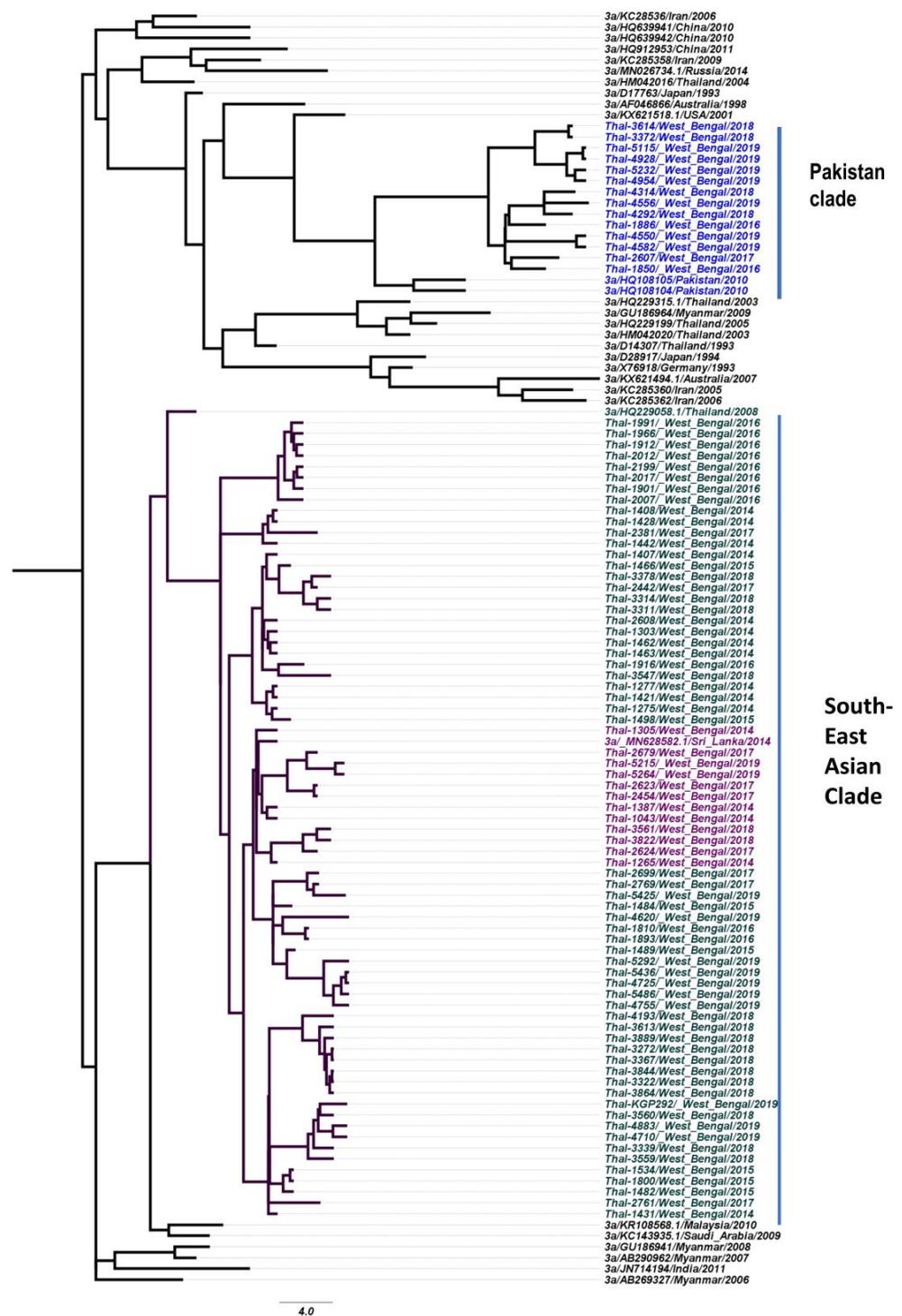
Figure 2. (3a Phylogeography) The HCV genotype 3a maximum clade tree was created using first 300 bp of the core genomic region (Respective of 342 nt to 642 nt of the strain H77-accession number NC-004102). The Bayesian MCMC BEAST software package version 1.10.4 was used to conduct the study. The tree was constructed using reference sequences from Iran: KC28536, KC285358, KC285360, KC285362; China: HQ639941, HQ639942, HQ912953; Russia: MN026734; Thailand: HM042016, HQ229199, HM042020, HQ229315, D14307, HQ229058; Japan: D17763, D28917; Australia: AF046866, KX621518; Pakistan: HQ108105, HQ108104; Myanmar: GU186964, GU186941, AB290962, AB269327; Saudi Arabia: KC143935; Malaysia: KR108568; Sri Lanka: MN628582; USA: KX621518; Germany: X76918; India: JN714194. The prefix "Thal" was used to encode local sequences.
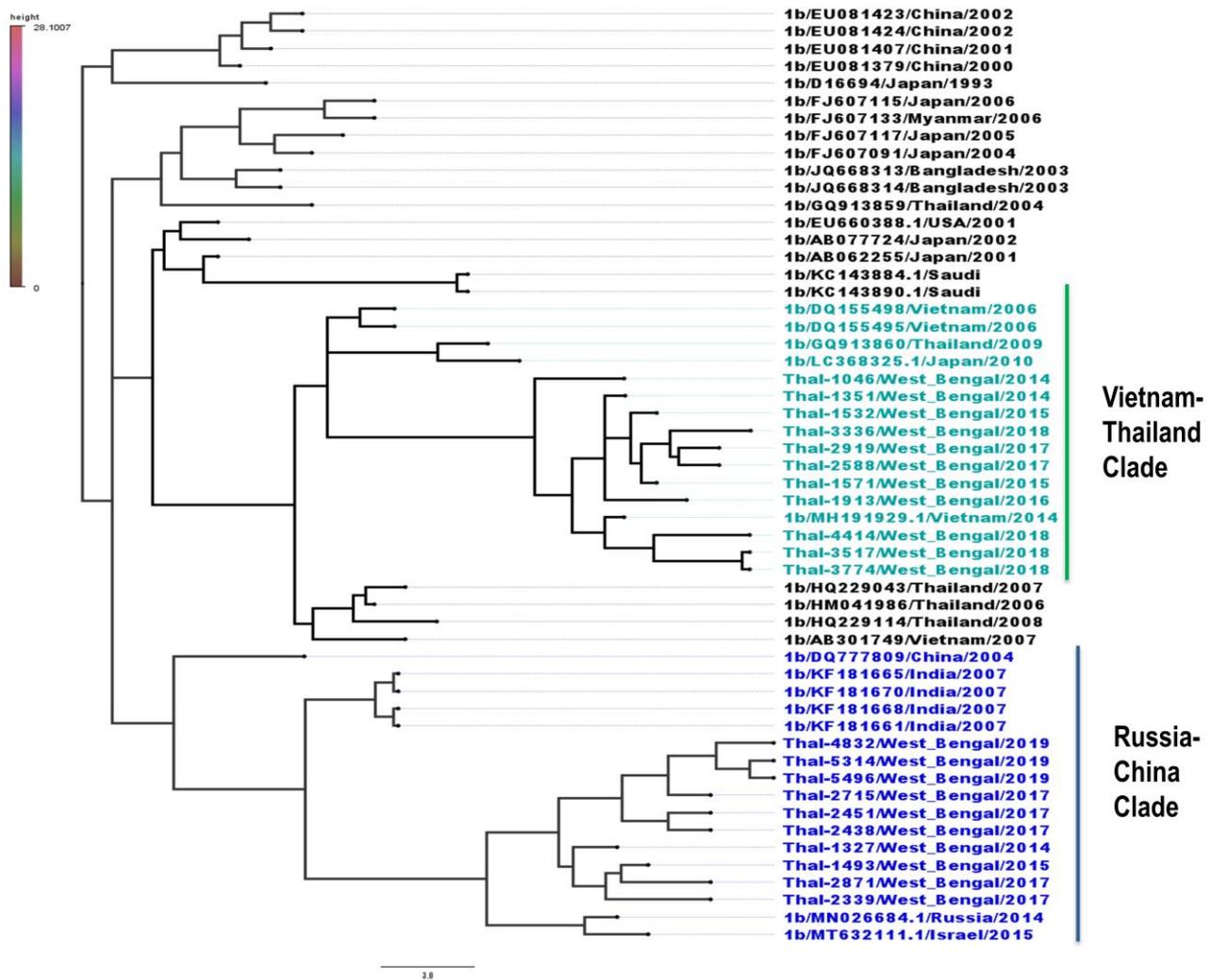
Figure 3. (1b Phylogeography) The HCV genotype 1b maximum clade tree was created using first 300 bp of the core genomic region (respective of 342 nt to 642 nt of the strain H77-accession number NC-004102). The Bayesian MCMC BEAST software package version 1.10.4 was used to conduct the study. The tree was constructed using reference sequences from China: EU081423, EU081424, EU081407, EU081379, DQ777809; Japan: AB077724, AB062255, D16694, FJ607115, FJ607117, FJ607117, FJ607091, LC368325; Thailand: GQ913859, HQ 229043, HQ229114, GQ913860; Vietnam: DQ155498, DQ155495, MH191929, AB301749; Bangladesh: JQ668313, JQ668314; Saudi Arabia: KC143884, KC143890; Myanmar: FJ607133; USA: EU660388; Russia: MN026684; Israel: MT632111; India: KF181665, KF181670, KF181668, KF181661. The prefix "Thal" was used to encode local sequences.

4. Discussion

HCV infection is a very common TTIs in multi-transfused β-thalassemic individuals causing serious health complications in this high-risk group population. The epidemiology and genotype distribution of HCV infection in thalassemic individuals is not very well reported from India or other countries, especially the HCV genotypes and geographic distribution. Determination of HCV genotypes and subtypes is essential to identify the source of HCV infection in a specified population and it is also required for clinical management, therapeutic intervention, and development of an effective HCV vaccine. This study was carried out over a period of six years (January, 2014 to December, 2019) in 917 HCV sero-reactive β-thalassemic patient samples that were collected from a total of 10 transfusion centers where patients received blood from their respective district TCs of West Bengal on a regular basis (Figure 1a). The prevalence of active HCV infection (state of viremia) in some of these transfusion centers was very high (Table 1). This research implies

that TCs could be hotspots for HCV transmission. It can also be postulated that due to improper screening of donor's blood during the seronegative "window" period, serological tests such as Tri-Dot assays and ELISA, which are more commonly used for donor blood screening in the blood banks, are insufficient to detect active HCV infection in donor blood. Hence, HCV-infected blood is being transfused to the thalassemic recipients. Therefore, the maintenance of strict transfusion protocols using more sensitive NAT-based RT-PCR screening methods is essential to check the donor's blood for probable infections before it is used for transfusion to prevent the transmission of HCV in this high-risk group population.

Data from this study revealed that β-thalassemic patients within the age group of 11–15 years had a higher percentage of HCV clearance (Table 3). From the findings, we concluded that the probability of HCV clearance is higher in teenage patients (11–15 years) as compared to adult patients (>25 years). This may be because of the increase in serum ferritin levels in older thalassemic patients due to repeated blood transfusions that facilitates HCV replication [25] and leads to faster progression to liver cirrhosis and HCC [26].

It has been previously reported that women can clear acute HCV infection at a higher rate than men [27]. Data from this study also provide evidence for the first time that in HCV-infected multi-transfused β-thalassemic Indian patients, thalassemic females were more likely to spontaneously clear the virus than thalassemic males (Table 3) and in addition the females had a lower chance of HCV RNA positivity than males (Table 1). This difference in the HCV clearance rate based on gender is still an unresolved issue but there are a few plausible explanations. In most viral infections, basal immune responses are higher in females than in males and low doses of estrogen have been suggested to modulate antiviral innate and adaptive immune responses [28]. Secondly, women have a slower rate of liver disease progression including liver cirrhosis and HCC [29]; thus, it is not surprising that the burden of HCV disease complications is lesser and the rate of infection clearance is higher in thalassemic females as compared to thalassemic males. In addition to these factors, several other yet to be determined factors such as menstrual cycles and genetic and immunological factors may play a key role in favoring HCV clearance in thalassemic females. It has been found that β-thalassemia patients had experienced menarche at 12–17 years depending on iron chelation therapy (ICT) [30] which somewhat justified our findings (Table 3) where it could be easily seen that female patients in the age group 11–15 years more efficiently cleared the infection. Studying the effects of the menstrual cycle in thalassemia-major women in India is challenging due to a lack of awareness about puberty and menstrual cycles in the Indian female population. Additionally, it is extremely difficult to collect data as women often hesitate to disclose such information publicly.

In this study, six HCV subtypes were noticed and they were 1a, 1b, 3a, 3b, 3g, and 3i (Figure 1b,c) of which genotypes 3g and 3i have not been previously reported from this region and this is the first time that these two HCV subtypes were found circulating in β-thalassemic patients of West Bengal. Another interesting finding is the higher predominance of HCV genotype 3 (especially subtype 3a) among β-thalassemic patients irrespective of the fact that the transfusion centers from where these patients received blood transfusions were all located in different districts. A similar pattern of HCV genotype 3 (subtype 3a) predominance among β-thalassemic patients in Iran [31] has been previously reported. HCV subtype 3a is mainly distributed through central, middle-east, and south-east Asian countries [32] which comprise the 'Global Thalassemia belt' [33]. In the South American sub-continent, thalassemia cases can be found in Brazil [34] and, surprisingly, HCV genotype 3a is present at a relatively high rate [35] there as well. Although the predominance of HCV genotype 3 (63.85%) in the Indian population has been reported earlier [36], our study reports that the prevalence of HCV genotype 3 among the Indian β-thalassemic teenage patients is even higher (87.79%). Several factors influence HCV replication, viz., host immunity, age, hepatocellular damage, and whether the infection is acute or chronic, and these factors might provide a replicative advantage to one HCV genotype in a specific patient population over other genotypes [37]. Although a detailed analysis is necessary to prove this from our observations it can be preliminarily speculated that HCV genotype

3a has a replication advantage over other HCV genotypes during the natural course of infection in the β-thalassemic Indian population. HCV genotype 3 is also associated to more aggressive hepatic steatosis and liver cancer [38–40], which may impart a serious life threat to β-thalassemia patients and further reduce their lifespan due to additional complications such as iron overloading, splenomegaly, etc.

Phylogeographic analysis of HCV strains assists in monitoring the virus for molecular tracing, ancestral studies, genetic assessments, clinical, and therapeutic intervention. Hence, phylogeographic analysis was performed with the two most commonly found HCV genotypes among β-thalassemia patients in West Bengal, 3a and 1b, to obtain information about the hierarchal relationships and genetic evolution. The data suggest that the majority of the HCV 3a isolates from thalassemia patients in West Bengal have sequence similarities with HCV strains from neighboring countries such as Pakistan, Sri Lanka, and Thailand (Figure 2) whereas the majority of the HCV 1b isolates of this region had sequence similarities to HCV strains of Vietnam, Thailand, Russia, and China (Figure 3), suggesting that the point of origin of the circulating HCV strains 3a and 1b can be traced back to the adjacent countries that share a border with India. However, we could not perform phylogeographic analyses of HCV strains 1a, 3i, and 3g as we did not isolate sufficient numbers of these strains from this study population to undertake any such analytical study.

Although this is the first detailed molecular epidemiological study of HCV infection among thalassemia patients in India, there were a few limitations. Firstly, 80% of the study population belonged to rural areas with poor economic backgrounds. There was a significant lack of awareness regarding thalassemia, how it spreads as well as the risk of Transfusion Transmitted Infections (TTIs). This study also included patients from districts known to be comprised of ethnic tribes [41,42]. Due to illiteracy and lack of awareness, a few of our patients lost their medical records and showed unwillingness to comply with medical professionals. These factors need to be addressed to further strengthen such studies. Furthermore, the NGS-based whole genome data of HCV could have enhanced the evolutionary aspect of this epidemiological study [43]. Hence, a combination of harm reduction measures, extensive sampling, and deep sequencing of HCV will help us better understand HCV pathogenesis and spread.

5. Conclusions

This study is one of the first multi-district, detailed, and thorough analyses of HCV viremia and genotypic distribution among β-thalassemia patients of Eastern India. Our study indicated that male thalassemia patients are more prone to chronic HCV infection than females, and those in rural areas and economically weaker sections are also at a higher risk of HCV infection. Thalassemia patients between the ages of 11–15 years and female β-thalassemic patients had a higher probability of spontaneously clearing the virus. Two major genotypes with six subtypes are circulating within this study population and the most prevalent HCV strains were 3a and 1b which share a sequence similarity with neighboring countries such as China, Pakistan, Myanmar, and Bangladesh, suggesting their probable migration into the Indian sub-continent. This investigation shows two new HCV subtypes (3g and 3i) which are not very common in this region. The prevalence of HCV infection is extremely high among Indian thalassemic patients. The introduction of more sensitive NAT-based tests for HCV detection in donor blood is an urgent requirement to overcome this situation.

References

1. Zeller, H.; Rabarijaona, L.; Rakoto-Andrianarivelo, M.; Boisier, P. Prevalence of hepatitis C virus infection in the general population of Madagascar. *Bull. Soc. Pathol. Exot.* **1997**, *90*, 3–5.
2. Smith, D.B.; Bukh, J.; Kuiken, C.; Muerhoff, A.S.; Rice, C.M.; Stapleton, J.T.; Simmonds, P. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: Updated criteria and genotype assignment web resource. *Hepatology* **2014**, *59*, 318–327. [CrossRef] [PubMed]
3. Tsai, S.-M.; Kao, J.-T.; Tsai, Y.-F. How hepatitis C patients manage the treatment process of pegylated interferon and ribavirin therapy: A qualitative study. *BMC Health Serv. Res.* **2016**, *16*, 247. [CrossRef] [PubMed]
4. Cavalcante, L.N.; Lyra, A.C. Predictive factors associated with hepatitis C antiviral therapy response. *World J. Hepatol.* **2015**, *7*, 1617–1631. [CrossRef] [PubMed]
5. Diagnosis & Management of Viral Hepatitis. Available online: https://www.inasl.org.in/diagnosis-management-viral-hepatitis.pdf (accessed on 15 June 2023).
6. Ghany, M.G.; Morgan, T.R.; Panel, A.H.C.G.; Marks, K.M.; Wyles, D.L.; Aronsohn, A.I.; Bhattacharya, D.; Broder, T.; Falade-Nwulia, O.O.; Feld, J.J.; et al. Hepatitis C Guidance 2019 Update: American Association for the Study of Liver Diseases–Infectious Diseases Society of America Recommendations for Testing, Managing, and Treating Hepatitis C Virus Infection. *Hepatology* **2020**, *71*, 686–721. [CrossRef]
7. Pozzetto, B.; Memmi, M.; Garraud, O.; Roblin, X.; Berthelot, P. Health care-associated hepatitis C virus infection. *World J. Gastroenterol.* **2014**, *20*, 17265–17278. [CrossRef] [PubMed]
8. Cao, A.; Galanello, R. Beta-thalassemia. *Anesthesia Analg.* **2010**, *12*, 61–76. [CrossRef]
9. Vidja, P.J.; Vachhani, J.H.; Sheikh, S.S.; Santwani, P.M. Blood Transfusion Transmitted Infections in Multiple Blood Transfused Patients of Beta Thalassaemia. *Indian J. Hematol. Blood Transfus.* **2011**, *27*, 65–69. [CrossRef]
10. Mohanty, D.; Colah, R.B.; Gorakshakar, A.C.; Patel, R.Z.; Master, D.C.; Mahanta, J.; Sharma, S.K.; Chaudhari, U.; Ghosh, M.; Das, S.; et al. Prevalence of β-thalassemia and other haemoglobinopathies in six cities in India: A multicentre study. *J. Community Genet.* **2013**, *4*, 33–42. [CrossRef]
11. Joseph, N.; Pai, S.; Sengupta, S.; Bharadwaj, S.; Dhawan, S.; Khare, K. A clinico-epidemiological study of thalassemia cases in India. *J. Nat. Sci. Biol. Med.* **2018**, *9*, 236. [CrossRef]
12. Colah, R.B.; Seth, T. Thalassemia in India. *Hemoglobin* **2022**, *46*, 20–26. [CrossRef]
13. Biceroglu, S.U.; Turhan, A. Probable hepatitis C virus transmission from a seronegative blood donor via cellular blood products. *Blood Transfus.* **2014**, *12*, s69–s70. [CrossRef]

14. Finianos, A.; Matar, C.F.; Taher, A. Hepatocellular Carcinoma in β-Thalassemia Patients: Review of the Literature with Molecular Insight into Liver Carcinogenesis. *Int. J. Mol. Sci.* **2018**, *19*, 4070. [CrossRef]

15. Gluba-Brzózka, A.; Franczyk, B.; Rysz-Górzyńska, M.; Rokicki, R.; Koziarska-Rościszewska, M.; Rysz, J. Pathomechanisms of Immunological Disturbances in β-Thalassemia. *Int. J. Mol. Sci.* **2021**, *22*, 9677. [CrossRef]

16. Petruzziello, A.; Marigliano, S.; Loquercio, G.; Cozzolino, A.; Cacciapuoti, C. Global epidemiology of hepatitis C virus infection: An up-date of the distribution and circulation of hepatitis C virus genotypes. *World J. Gastroenterol.* **2016**, *22*, 7824–7840. [CrossRef]

17. Satsangi, S.; Chawla, Y.K. Viral hepatitis: Indian scenario. *Med. J. Armed Forces India* **2016**, *72*, 204–210. [CrossRef] [PubMed]

18. Biswas, A.; Sarkar, K.; Firdaus, R.; Saha, K.; Gupta, D.; Ghosh, M.; Chowdhury, P.; Bhattacharyya, D.; Bhattacharyya, M.; Sadhukhan, P.C. Prevalence of ANTI-HCV, HBSAG, HIV among multi-transfused thalassemic individuals and their socio-economic background in Eastern India. *Asian J. Pharm. Clin. Res.* **2016**, *9*, 290–294.

19. Saha, K.; Firdaus, R.; Biswas, A.; Mukherjee, A.; Sadhukhan, P. A novel nested reverse-transcriptase polymerase chain reaction method for rapid hepatitis C virus detection and genotyping. *Indian J. Med. Microbiol.* **2014**, *32*, 130–136. [CrossRef] [PubMed]

20. Firdaus, R.; Biswas, A.; Saha, K.; Mukherjee, A.; Chaudhuri, S.; Chandra, A.; Konar, A.; Sadhukhan, P.C. Impact of Host IL28B rs12979860, rs8099917 in Interferon Responsiveness and Advanced Liver Disease in Chronic Genotype 3 Hepatitis C Patients. *PLoS ONE* **2014**, *9*, e99126. [CrossRef]

21. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef] [PubMed]

22. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [CrossRef] [PubMed]

23. Drummond, A.J.; Ho, S.Y.W.; Phillips, M.J.; Rambaut, A. Relaxed Phylogenetics and Dating with Confidence. *PLoS Biol.* **2006**, *4*, e88. [CrossRef]

24. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef] [PubMed]

25. Zou, D.-M.; Sun, W.-L. Relationship between Hepatitis C Virus Infection and Iron Overload. *Chin. Med. J.* **2017**, *130*, 866–871. [CrossRef] [PubMed]

26. Marcus, E.L.; Tur-Kaspa, R. Chronic Hepatatis C Virus Infection in Older Adults. *Clin. Infect. Dis.* **2013**, *41*, 1606–1612.

27. Guy, J.; Peters, M.G. Liver disease in women: The influence of gender on epidemiology, natural history, and patient outcomes. *Gastroenterol. Hepatol.* **2013**, *9*, 633–639.

28. Ruggieri, A.; Gagliardi, M.C.; Anticoli, S. Sex-Dependent Outcome of Hepatitis B and C Viruses Infections: Synergy of Sex Hormones and Immune Responses? *Front. Immunol.* **2018**, *9*, 2302. [CrossRef]

29. Baden, R.; Rockstroh, J.K.; Buti, M. Natural History and Management of Hepatitis C: Does Sex Play a Role? *J. Infect. Dis.* **2014**, *209* (Suppl. 3), S81–S85. [CrossRef]

30. De Sanctis, V.; Di Maio, S. A Retrospective Long-Term Study on Age at Menarche and Menstrual Characteristics in 85 Young Women with Transfusion-Dependent β-Thalassemia (tdt) Born between 1965 and 1995. *Mediterr. J. Hematol. Infect. Dis.* **2021**, *13*, e2021040. [CrossRef]

31. Shahraki, T.; Shahraki, M.; Moghaddam, E.S.; Najafi, M.; Bahari, A. Determination of Hepatitis C Genotypes and the Viral Titer Distribution in Children and Adolescents with Major Thalassemia. *Iran. J. Pediatr.* **2010**, *20*, 75–81.

32. Gower, E.; Estes, C.; Blach, S.; Razavi-Shearer, K.; Razavi, H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J. Hepatol.* **2014**, *61*, S45–S57. [CrossRef]

33. De Sanctis, V. β-thalassemia distribution in the old world: A historical standpoint of an ancient disease. *Mediterr. J. Hematol. Infect. Dis.* **2017**, *9*, e2017018. [CrossRef] [PubMed]

34. Bonini-Domingos, C.R. Thalassemia screening in Brazil: Results for 20 years. *Rev. Bras. Hematol. Hemoter.* **2004**, *26*, 288–289. [CrossRef]

35. Paraná, R.; Vitvitski, L.; Berby, F.; Portugal, M.; Cotrim, H.P.; Cavalcante, A.; Lyra, L.; Trepo, C. HCV infection in northeastern Brazil: Unexpected high prevalence of genotype 3a and absence of African genotypes. *Arq. de Gastroenterol.* **2000**, *37*, 213–216. [CrossRef] [PubMed]

36. Christdas, J.; Sivakumar, J.; David, J.; Daniel, H.; Raghuraman, S.; Abraham, P. Genotypes of hepatitis C virus in the Indian sub-continent: A decade-long experience from a tertiary care hospital in South India. *Indian J. Med. Microbiol.* **2013**, *31*, 349–353. [CrossRef]

37. Yang, Z.; Yu, Y.; Zhang, H.; Shang, G.; Gao, J.; Jiang, J.-D.; Peng, Z. Replication priority of hepatitis C virus genotype 2a in a Chinese cohort. *Acta Pharm. Sin. B* **2014**, *4*, 266–269. [CrossRef] [PubMed]

38. Westin, J.; Nordlinder, H.; Lagging, M.; Norkrans, G.; Wejstål, R. Steatosis accelerates fibrosis development over time in hepatitis C virus genotype 3 infected patients. *J. Hepatol.* **2002**, *37*, 837–842. [CrossRef]

39. Nkontchou, G.; Ziol, M.; Aout, M.; Lhabadie, M.; Baazia, Y.; Mahmoudi, A.; Roulot, D.; Ganne-Carrié, N.; Grando-Lemaire, V.; Trinchet, J.-C.; et al. HCV genotype 3 is associated with a higher hepatocellular carcinoma incidence in patients with ongoing viral C cirrhosis. *J. Viral Hepat.* **2011**, *18*, e516–e522. [CrossRef]

40. Bochud, P.-Y.; Cai, T.; Overbeck, K.; Bochud, M.; Dufour, J.-F.; Müllhaupt, B.; Borovicka, J.; Heim, M.; Moradpour, D.; Cerny, A.; et al. Genotype 3 is associated with accelerated fibrosis progression in chronic hepatitis C. *J. Hepatol.* **2009**, *51*, 655–666. [CrossRef]

41. Mukhopadhyay, D.K.; Mukhopadhyay, S.; Biswas, A.B. Enduring starvation in silent Population: A study on prevalence and factors contributing to household food security in the tribal population in Bankura, West Bengal. *Indian J. Public Health* **2010**, *54*, 92–97. [CrossRef]

42. Modak, B.K.; Gorai, P.; Dhan, R.; Mukherjee, A.; Dey, A. Tradition in treating taboo: Folkloric medicinal wisdom of the aboriginals of Purulia district, West Bengal, India against sexual, gynaecological and related disorders. *J. Ethnopharmacol.* **2015**, *169*, 370–386. [CrossRef] [PubMed]

43. Lu, L.; Li, C.; Yuan, J.; Lu, T.; Okamoto, H.; Murphy, D.G. Full-length genome sequences of five hepatitis C virus isolates representing subtypes 3g, 3h, 3i and 3k, and a unique genotype 3 variant. *J. Gen. Virol.* **2013**, *94*, 543–548. [CrossRef] [PubMed]