TITLE PAGE

A short and informative title:

Study of Hepatitis C Virus infection among multi-transfused patients with inherited β-globin synthesis gene defect, in the eastern region of India

A short running title:

HCV infection among multi-transfused β -thalassemia individuals

The full name(s) of the author(s):

Supradip Dutta[1*], Aritra Biswas[1*], Promisree Choudhury[1*], Sagnik Bakshi, Prosanto Chowdhury[2], Maitreyee Bhattacharyya[3], Sharmistha Chakraborty[4], Shanta Dutta[1], Provash C Sadhukhan[1]

[*] ***Authors contributed equally***

The affiliation(s) and address(es) of the author(s):

[1] Indian Council of Medical Research-National Institute of Cholera and Enteric Diseases
P-33, Scheme XM, CIT Road, Beliaghata, Kolkata-700010, West Bengal, India.

[2] Peerless Hospital & B. K. Roy Research Centre
360 Panchasayar**,** Kolkata – 700094, West Bengal, India.

[3] Medical College Kolkata
88, College St, College Square, Kolkata- 700073, West Bengal, India.

[4] Burdwan Medical College and Hospital
Baburbag, Rajbati, Burdwan – 713104, West Bengal, India.

The e-mail address, and telephone number(s) of the corresponding author:

Provash Chandra Sadhukhan,
Email ID: sadhukhan.pc@icmr.gov.in
Tel: 91-9830546338

Disclosure of potential conflicts of interest*:*

The authors declare they have no conflict of interest.

Funding:

This study was funded by the Indian Council of Medical Research (ICMR), India and Department of Science and Technology (DST), Government of West Bengal, India.

Abstract:

Background: Post transfusion acquired HCV infection is common in high-risk group individuals such as multi-transfused β-thalassemia patients who depend on regular blood transfusions. This study was conducted to determine epidemiology and distribution 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-thalassemic patients and processed for viral RNA isolation followed by nested RT-PCR for qualitative viremia detection. HCV genotype was determined by amplifying partial HCV core gene by nested RT-PCR, DNA sequencing and using NCBI genotyping tools. Phylogenetic and phylogeographic studies were performed with online MEGA-X and BEAST 1.10.0 software respectively.

Results: Out of 917 multi-transfused HCV sero-reactive β-thalassemic patients, 598 (65.21%) were positive for HCV RNA while 250 (41.80%) had spontaneously cleared the virus. Female thalassemic patients and individuals belonging to ages 10-14 years had higher chances of spontaneous clearance. The most prevalent circulatory HCV genotype was 3a (78.11%) followed by 1b (12.20%). Phylogeographic analyses revealed that the 3a strains share similarity with Pakistan, Sri Lanka and Thailand whereas the 1b strains share similarity with Thailand, Vietnam, Russia and China.

Conclusion: The prevalence of HCV infection is very high among Indian β-thalassemic patients, necessitates a critical look into the prevailing transfusion practices and requires implementation of more rigid donor screening criteria to decrease the rate of transfusion transmitted HCV infection,

especially in multi-transfused thalassemic patients. The use of more sensitive NAT based assays for HCV detection in donor blood is a compressing need of the hour.

Keywords:

Hepatitis C virus, Genotype distribution, Thalassemia, West Bengal, India

INTRODUCTION

HCV is an enveloped, positive sense, ~ 9.6 kb ssRNA virus belong to Hepacivirus genus and *Flaviviridae* family that can cause chronic liver diseases, including cirrhosis of liver and hepatocellular carcinoma (HCC). HCV mainly transmits through blood, blood products and body fluids infecting an estimated 200 million people globally [1]. After leaving the bloodstream, HCV enters the liver and replicates in hepatocytes, resulting acute or chronic liver infection. Based on viral genetic diversity, HCV has been classified into seven genotypes (1-7) and over 67 subtypes [2]. It is extremely important to know the HCV genotype of an infected patient for genotype specific Direct Acting Antivirals (DAAs) treatment [3]. HCV genotypes differ from each other not only in nucleotide sequences but also in geographical distribution [4].

β-thalassemia is an inherited autosomal recessive disorder arises from reduced or lack of synthesis of the beta globin chain of hemoglobin [5]. 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 [6]. Approximately one-tenth of the world's thalassemic patients are born in India every year of which approximately 1-3% in southern India and 3-15% in northern India are carriers [7]. Since 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 [8].

Genomic diversity of HCV in thalassemia patients depends on geographical localization. Different countries may have different genotype distribution[9]. In India, genotype 3 is prevalent in the north, east and west and genotype 1 in the south[10]. Although, a hospital-based study on the

prevalence and genotypic distribution of HCV in thalassemia patients were conducted [11], 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 multi-transfused β-thalassemia patient population in the Indian state of West Bengal from 2014-2019.

MATERIALS AND METHODS

Ethical Statement: Written informed consent was obtained from the patients before including them for this study. This study protocol complied with the Helsinki Declaration of 1975, amended in 2013, and was approved by the Institutional Ethical Committee, Indian Council of Medical Research-National Institute of Cholera and Enteric Diseases (ICMR-NICED), Kolkata, India.

Study Design: A total of 917 HCV sero-reactive β -thalassemic patients were enrolled in this study. Patient's blood samples were collected by venipuncture in clot vials followed by 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 as TC 1-10 and are located around eight southern districts namely, Bankura, East Burdwan, West Burdwan, East Midnapur, West Midnapur, Howrah, Hooghly and Kolkata in West Bengal, the eastern state of India. All laboratory works were done at ICMR-NICED, Kolkata. HCV sero-reactivity were rechecked by HCV ELISA (HCV Mono ELISA, BioRad). HCV RNA positive patients were classified into seven age groups (AG), 1 to 7, depending on their age namely, AG-1 (2-5 years), AG-2 (6-9 years), AG-3 (10-14 years), AG-4 (15-19 years), AG-5 (20-24 years), AG-6 (25-29 years) and AG-7 (30 years and above). Patient age ranged from 2 to 35 years with a mean (±SD) of 11.5 ± 7.5 years. Spontaneous clearance was defined as undetectable level of HCV RNA (viral

load < 50 IU/ml) by quantitative Real Time PCR in two consecutive testing in 6 months apart. Patients who achieved spontaneous clearance in this study were followed up to 2 years after initial detection of infection.

Detection of HCV RNA and genotyping: Viral RNA was isolated from all sero-reactive serum samples using QIAamp viral RNA mini kit (Qiagen, Hilden, Germany) according to manufacturer's instructions. Briefly, viral RNA was extracted from 140µl serum, eluted in 50µl elution buffer and stored at -80ºC for further use.

Detection of HCV viral RNA was done by nested RT-PCR based on 5´UTR of HCV genome as described previously [12]. Briefly, first round one tube RT-PCR was done in 20µl total reaction volume containing 2µl isolated RNA and second round nested PCR was performed in 25µl total volume using 2µl 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, USA). Quantitative HCV RNA was estimated using a Qiagen quantitative real time RT-PCR (qRT-PCR) kit (QuantiFast Pathogen RT-PCR plus IC Kit). The HCV primer and probe sequences were directed against the 5´ UTR of the HCV genome[13] using the 4th WHO International Standard for HCV (NIBSC code 06/102) as standard.

For the determination of HCV genotype, nested RT-PCR amplified amplicons of partial HCV core gene (405 bp) [14] were gel purified and directly used for DNA sequencing analyses in an automated DNA sequencer, model 3130XL (ABI, USA) using Big Dye terminator 3.1 kit (Applied Biosystems, USA). HCV genotyping were determined by NCBI genotyping tool (https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi) .

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 86 laboratory isolated 3a strains (Accession Numbers, MN635186-MN635242, MW330397-330425), twenty one 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 (Accession number- MW368886) and 1a strains (Accession number-MW368892) were aligned with HCV reference strains using Molecular Evolutionary Genetics Analysis tool (MEGA-X)[15]. The evolutionary history was inferred using the Neighbor-Joining (NJ) method[16]. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-3826.33) was shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+*G*, parameter = 0.4698)).

Phylogeographic analysis: To understand the phylogeographic relationship among neighboring countries and HCV genotype distribution with evolutionary linkages in β -thalassemia population, phylogeographic trees were constructed using Bayesian Evolutionary Analysis Sampling Trees (BEAST), package 1.10.4 (http://beast.community/programs). HCV core sequences from the following countries, India, China, Pakistan, Bangladesh, Myanmar, Thailand, Japan, Vietnam,

Iran, Australia were downloaded from ViPR (Virus Pathogen Resource) database (www.viprbrc.org/brc/home.spg?decorator=vipr), HCV sequence database (https://hcv.lanl.gov/content/index) and NCBI. First 300 bp (Respective of 371 nt to 671 nt of the 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 clocks setting, uncorrelated relaxed clock type was used [17] along with the Bayesian Skyline Coalescent model. It was also ensured achievement of Effective Sampling Sizes (ESS) >200 in TRACER (http://beast.community/tracer)[18]. The result of BEAST analyses was evaluated in TRACER. Tree Annotator (https://beast.community/treeannotator) and Fig-Tree (https://beast.community/figtree) software were used to construct the final tree.

RESULTS

Distribution of HCV viremia in ten transfusion centers:

A total of 917 HCV sero-reactive serum samples were collected over six years from multi-transfused β-globulin synthesis defective thalassemia patients at either of the 10 TCs located around ten districts of West Bengal, India as indicated in **Figure 1**. Overall, HCV RNA was detected in 65.21% (n=598) sero-reactive samples by either qualitative nested RT-PCR or quantitative real time RT-PCR. Prevalence of HCV RNA as observed in the 10 TCs were 81.25% (TC-1, n=30), 70.27% (TC-2, n=30), 70.58% (TC-3, n=28), 64.20% (TC-4, n=213), 58.51% (TC-5, n=65), 56.75% (TC-6, n=24), 82.27% (TC-7, n=75), 60.52% (TC-8, n=27), 52.56% (TC-9, n=47) and 62.33% (TC-10, n=59) respectively. The distribution of HCV viremia among the different TCs is depicted in **Table 1**.

HCV genotype distribution, demographic and risk factors:

598 partial core gene sequences (405 bp) were amplified and sequenced for HCV genotyping and phylogenetic analyses. Sequences were aligned with references sequences by 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. 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 (78.11%, n = 467) followed by subtype 1b (12.03%, n=72) and 3b (8.69%, n=52), while subtypes 3g, 3i and 1a were found in only 3 (0.50%), 2 (0.33%) and 1 (0.17%) patient(s) respectively. Studies on the distribution of HCV genotypes in each TC individually revealed that HCV genotype 3a was detected in all the 10 TCs (TC1-TC10) and was the most prevalent HCV strain among all these 10 centers. HCV strains 1b and 3b could be found

in only seven (TC-1 to TC-7) and nine (TC-1 to TC-9) TCs respectively, whereas HCV strains 1a, 3g and 3i were detected from one (TC-6), two (TC-5, TC-9) and one (TC-7) TC(s) respectively. Population based DNA sequencing data from the partial core region of 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 gene bank with accession numbers as mentioned in the materials and methods section.

Phylogeography:

To examine 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 majority of HCV genotype 3a sequences isolated in our laboratory formed three distinct clusters resulting in the Pakistan clade, Thailand clade and 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)**.

DISCUSSION

HCV infection is one of the most 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 HCV genotypes and geographic distribution. Determination of HCV genotypes and subtypes is essential to know 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 districts of West Bengal on regular **(Figure 1b)**. The prevalence of active HCV infection (state of viremia) in some of these transfusion centers was as high as 82% **(Table 1)**. This data suggests that TCs may be act as hotspots for transmitting HCV infection. 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 not sufficient to detect the presence of active HCV infection in donor blood. Hence, HCV infected blood is getting transfused to the thalassemic recipients. Therefore, maintenance of strict transfusion protocols using more sensitive NAT based RT-PCR screening methods are essential to check 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 6-9 years had a higher prevalence of HCV viremia **(Table 1)** which is very rarely reported and age group of 10-14 years had a higher percentage of HCV clearance **(Table 2)**. We also found that patients with lower age (6-14 years) had higher HCV clearance as compared to patients with higher age (20-29 years) suggesting that low age thalassemic patients have a higher probability of clearing the infection than older patients. This could be because of the increase in serum ferritin levels in older thalassemic patients due to repeated blood transfusions that facilitates HCV replication[19] and leads to faster progression to liver cirrhosis and HCC[20].

It has been previously reported that women can clear acute HCV infection at a higher rate than men[21]. Data from this study also provides 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 2)**. This difference in 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[22]. Secondly, women have a slower rate of liver disease progression including liver cirrhosis and HCC[23], thus it is not surprising that the burden of HCV disease complications and the rate of infection clearance is lower in thalassemic females as compared to thalassemic males. In addition to these factors, several other yet to be determined like, genetic and immunologic factors may play a key role in favoring HCV clearance in the thalassemic females.

In this study, six HCV subtypes were noticed and they were 1a, 1b, 3a, 3b, 3g and 3i **(Figure 1a, 1c)** 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 (specially 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. We had also observed similar pattern of HCV genotype 3 (subtype 3a) predominance among β-thalassemic patients in Iran[24]. It was also found that HCV subtype 3a mainly distributed through central, middle-east, and south-east Asian countries[25-30]when almost all these countries were situated on 'Global Thalassemia belt'[31] . In South American sub-continent thalassemia cases can be found in Brazil[32] and with more surprising observation that HCV genotype 3a was also present

over there in a relatively high rate[33]. Although the predominance of HCV genotype 3 (63.85%) in the Indian population have been reported earlier[34] but our study reports that the prevalence of HCV genotype 3 among the Indian β-thalassemic patients is even higher (87.79%). Several factors are known to influence HCV replication such as host immunity, age, hepatocellular damage, whether the infection is acute or chronic and these factors might provide a replicative advantage to one HCV genotype in a specific patient population than other genotypes[35] . Although a more detailed analyses is necessary to prove this but from our observations it can be preliminarily speculated that HCV genotype 3a has a replication priority over other HCV genotypes during the natural course of infection in the β-thalassemic Indian population. HCV genotype 3 also associated to more aggressive hepatic steatosis and liver cancer[36–38] , which may impart a serious life threat to β-thalassemia patients and further reduce their lifespan due to additional complications like iron overloading, splenomegaly, etc.

Phylogeographic analysis of HCV strains assist in monitoring the virus for molecular tracing, ancestral studies, genetic assessments, clinical and therapeutic intervention. Hence, phylogeographic analysis was performed with 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. Data suggests that majority of the HCV 3a isolates from thalassemia patients in West Bengal have sequence similarities with HCV strains from neighboring countries like Pakistan, Sri Lanka and Thailand **(Figure 2)** whereas 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 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.

To summarize, this study provides the first multi-district, comprehensive and detail analyses of HCV viremia and genomic diversity among β-thalassemia patients in eastern part of India. Two major genotypes with 6 subtypes are circulating within this study population and the most prevalent HCV strains were 3a and 1b. This study provides the evidence of the presence of two new additional HCV strains 3i and 3g, which are not commonly reported from this region. Our study also indicated that young adults/ youth between the ages of 10-14 years and female β-thalassemic patients had a higher probability of spontaneously clearing the virus. Additionally, this study is also reported that laboratory isolated HCV strains 3a and 1b have sequence similarity with neighboring countries like, China, Pakistan, Myanmar and Bangladesh and may be those HCV strains migrated to this region.

Beta-Thalassemia patients are at very high risk for HCV infection due to repeated blood transfusion. Yet, very little studies have been reported about prevalence and genotype distribution of HCV in this neglected population which is the main focus of our study.

We have found very High HCV prevalence among beta-thalassemia patients specially in children of eastern part of India. They are also at a high risk of Cirrhosis because genotype 3 (specially 3a) is the most prevalent genotype in this population. Two new subtypes 3g and 3i also have been seen in this population which can be a future problem for genotype specific HCV Directly Acting Antivirals (DAAs) treatment and clinical management for thalassemia population in this region. Two major genotypes (3a and 1b) share similarities with neighbouring countries HCV strains.

When all World is fighting to eradicate HCV within 2030, Thalassemia population can be a major hurdle in that path, if proper clinical management and NAT based detections of donor blood are not implemented as early as possible.

Acknowledgements: We would like to acknowledge all the past and present members of Dr. Sadhukhan laboratory: Kallol Saha, Rushna Firdaus, Upasana Baskey, Priya K. Verma, Raina Das, Ronita De, Sudipta Biswas and Maya Halder for all their help with this study. We would like to thank the patients and their families for their kind cooperation and participation to perform this study. We would also like to thank our collaborative institutes for their help and support. We would also like to acknowledge the financial help from Indian Council of Medical Research, New Delhi, Govt. of India; Department of Science and Technology, Govt. of West Bengal for their financial support to carry out this study.

REFERENCES:

1. Zeller H, Rabarijaona L, Rakoto-Andrianarivelo M, et al. [Prevalence of hepatitis C virus infection in the general population of Madagascar]. *BullSocPatholExot*. 2017;90(1):3-5.

2. Smith DB, Bukh J, Kuiken C, et al. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: Updated criteria and genotype assignment web resource. *Hepatology*. 2014;59(1):318-327. doi:10.1002/hep.26744

3. Cavalcante LN, Lyra AC. Predictive factors associated with hepatitis C antiviral therapy response. *World J Hepatol*. 2015;7(12):1617-1631. doi:10.4254/wjh.v7.i12.1617

4. Pozzetto B, Memmi M, Garraud O, et al. Health care-associated hepatitis C virus infection. *World J Gastroenterol*. 2014;20(46):17265-17278. doi:10.3748/wjg.v20.i46.17265

5. Cao A, Galanello R. Beta-thalassemia. 2010. doi:10.1097/GIM.0b013e3181cd68ed

6. Vidja PJ, Vachhani JH, Sheikh SS, et al. Blood Transfusion Transmitted Infections in Multiple Blood Transfused Patients of Beta Thalassaemia. *Indian J Hematol Blood Transfus*. 2011;27(2):65-69. doi:10.1007/s12288-011-0057-3

7. Joseph N, Pai S, Sengupta S, et al. A clinico-epidemiological study of thalassemia cases in India. *J Nat Sci Biol Med*. 2018;9(2):236. doi:10.4103/jnsbm.JNSBM_224_17

8. Biceroglu SU, Turhan A, Doskaya AD, et al. Probable hepatitis C virus transmission from a seronegative blood donor via cellular blood products. *Blood Transfus*. 2014;12(SUPPL.1):69-70. doi:10.2450/2013.0285-12

9. Petruzziello A, Marigliano S, Loquercio G, et al. 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(34):7824-7840. doi:10.3748/wjg.v22.i34.7824

10. Satsangi S, Chawla YK. Viral hepatitis: Indian scenario. *Med J Armed Forces India*. 2016;72(3):204-210. doi:10.1016/j.mjafi.2016.06.011

11. Biswas A, Sarkar K, Firdaus R, et al. 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(1):290-294.

12. Saha K, Biswas A, Firdaus R, et al. A novel nested RT-PCR method for Rapid HCV Detection and Genotyping. 2014;(APRIL):1-12.

13. Firdaus R, Biswas A, Saha K, et al. Impact of Host IL28B rs12979860, rs8099917 in Interferon Responsiveness and Advanced Liver Disease in Chronic Genotype 3 Hepatitis C Patients. doi:10.1371/journal.pone.0099126

14. Sadhukhan P, Saha K, Firdaus R, et al. A novel nested reverse-transcriptase polymerase chain reaction method for rapid hepatitis C virus detection and genotyping. *Indian J Med Microbiol*. 2014;32(2):130. doi:10.4103/0255-0857.129782

15. Kumar S, Stecher G, Li M, Knyaz C, et al. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547-1549. doi:10.1093/molbev/msy096

16. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci*. 2004;101(30):11030-11035. doi:10.1073/pnas.0404206101

17. Drummond AJ, Ho SYW, Phillips MJ, et al. Relaxed Phylogenetics and Dating with Confidence. 2006;4(5). doi:10.1371/journal.pbio.0040088

18. Rambaut A, Drummond AJ, Xie D, et al. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol*. 2018;67(5):901-904. doi:10.1093/sysbio/syy032

19. Zou DM, Sun WL. Relationship between hepatitis C virus infection and iron overload. *Chin Med J (Engl)*. 2017;130(7):866-871. doi:10.4103/0366-6999.202737

20. Marcus EL, Tur-Kaspa R. Chronic hepatitis C virus infection in older adults. *Clin Infect Dis*. 2005;41(11):1606-1612. doi:10.1086/497597

21. Guy J, Peters MG. Liver disease in women: The influence of gender on epidemiology, natural history, and patient outcomes. *Gastroenterol Hepatol*. 2013;9(10):633-639.

22. Ruggieri A, Gagliardi MC, Anticoli S. Sex-dependent outcome of hepatitis B and C Viruses infections: Synergy of sex hormones and immune responses? *Front Immunol*. 2018;9(OCT):1-7. doi:10.3389/fimmu.2018.02302

23. Baden R, Rockstroh JK, Buti M. Natural history and management of hepatitis C: Does sex play a role? *J Infect Dis*. 2014;209(SUPPL. 3):S81-S85. doi:10.1093/infdis/jiu057

24. Shahraki T, Shahraki M, Moghaddam ES, et al. Determination of hepatitis C genotypes and the viral titer distribution in children and adolescents with major thalassemia. *Iran J Pediatr*. 2010;20(1):75-81. /pmc/articles/PMC3445996/?report=abstract. Accessed December 21, 2020.

25. Gower E, Estes C, Blach S, et al. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J Hepatol*. 2014;61(1):S45-S57. doi:10.1016/j.jhep.2014.07.027

26. Chan A, Patel K, Naggie S. Genotype 3 Infection: The Last Stand of Hepatitis C Virus. *Drugs*. 2017;77(2):131-144. doi:10.1007/s40265-016-0685-x

27. Stamouli M, Panagiotou I, Kairis D, et al. Genotype distribution in chronic hepatitis C patients in Greece. *Clin Lab*. 2012;58(1-2):173-176. doi:10.1016/j.ijid.2008.05.1097

28. Umer M. Hepatitis C virus prevalence and genotype distribution in Pakistan: Comprehensive review of recent data. *World J Gastroenterol*. 2016;22(4):1684. doi:10.3748/wjg.v22.i4.1684

29. Janahi EM, Al-Mannai M, Singh H, et al. Distribution of hepatitis C virus genotypes in bahrain. *Hepat Mon*. 2015;15(12). doi:10.5812/hepatmon.30300

30. Wasitthankasem R, Vongpunsawad S, Siripon N, et al. Genotypic Distribution of Hepatitis C Virus in Thailand and Southeast Asia. Hotta H, ed. *PLoS One*. 2015;10(5):e0126764. doi:10.1371/journal.pone.0126764

31. Sanctis V De, Kattamis C, Canatan D, et al. β-thalassemia distribution in the old world: An ancient disease seen from a historical standpoint. *Mediterr J Hematol Infect Dis*. 2017;9(1):2017018. doi:10.4084/mjhid.2017.018

32. Bonini-Domingos CR. Thalassemia screening in Brazil: results for 20 years. *Rev Bras Hematol Hemoter*. 2004;26(4):288-289. doi:10.1590/s1516-84842004000400011

33. Paraná R, Vitvitski L, Berby F, et al. HCV infection in northeastern Brazil: Unexpected high prevalence of genotype 3a and absence of African genotypes. *Arq Gastroenterol*. 2000;37(4):213-216. doi:10.1590/S0004-28032000000400005

34. Abraham P, Sivakumar J, Raghuraman S, et al. 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(4):349. doi:10.4103/0255-0857.118875

35. Yang Z, Yu Y, Zhang H, et al. Replication priority of hepatitis C virus genotype 2a in a Chinese cohort. *Acta Pharm Sin B*. 2014;4(4):266-269. doi:10.1016/j.apsb.2014.06.011

36. Westin J, Nordlinder H, Lagging M, et al. Steatosis accelerates fibrosis development over time in hepatitis C virus genotype 3 infected patients. *J Hepatol*. 2002;37(6):837-842. doi:10.1016/S0168-8278(02)00299-4

37. Nkontchou G, Ziol M, Aout M, 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(10). doi:10.1111/j.1365-2893.2011.01441.x

38. Bochud PY, Cai T, Overbeck K, et al. Genotype 3 is associated with accelerated fibrosis progression in chronic hepatitis C. *J Hepatol*. 2009;51(4):655-666. doi:10.1016/j.jhep.2009.05.016

Tables and Figures: Table 1

Variable	HCV genotypes		HCV positive
	Genotype 1 (n= 73)	Genotype 3 (n=525)	
Gender (Male/Female)	50/23(68.50%/31.50%)	331/194(63.04%/36.96%)	598
Age			
▪ 2-5	5(3.60%)	134(96.40%)	139
▪ 6-9	5(2.60%)	187(97.40%)	192
▪ 10-14	33(18.44%)	146(81.56%)	179
▪ 15- 19	14(29.16%)	34(70.84%)	48
▪ 20- 24	10(45.45%)	12(54.55%)	22
▪ 25- 29	2(22.22%)	7(77.78%)	9
▪ >30	4(44.44%)	5(55.56%)	9
Transfusion Centre (TC)			
▪ TC 1	4(13.33%)	26(86.67%)	30
▪ TC 2	3(10%)	27(90%)	30
▪ TC 3	4(14.28%)	24(85.72%)	28
▪ TC 4	28(13.14%)	185(86.86%)	213
▪ TC 5	8(12.30%)	57(87.70%)	65
▪ TC 6	3(12.5%)	21(87.5%)	24
▪ TC 7	4(5.34%)	71(94.66%)	75
▪ TC 8	4(14.81%)	23(85.19%)	27
▪ TC 9	6(12.76%)	41(87.24%)	47
▪ TC 10	9(15.25%)	50(84.75%)	59
Locality of Residence			
▪ Rural	54(11%)	437 (89%)	491
▪ Urban	19(17.76%)	88(82.24%)	107
Economic Class			
▪ Below poverty	62(12.20%)	446(87.80%)	508
▪ Above poverty	11((12.23%)	79(87.77%)	90

Table 2

	Age 2yr-5yr	Age 6yr-9yr	Age 10yr-14yr	Age 15yr-19yr	Age 20yr-24yr	Total number of Spontaneous Clearance
Male (total no. of follow-ups, n= 309 out of 381 patients)	30 (22.06%)	45 (33.09%)	52 (38.24%)	6 (4.41%)	3 (2.20%)	136 (44%)
Female (total no. of follow-ups, n= 176 out of 217 patients)	25 (21.93%)	39 (34.22%)	43 (37.72%)	5 (4.38%)	2 (1.75%)	114 (65%)

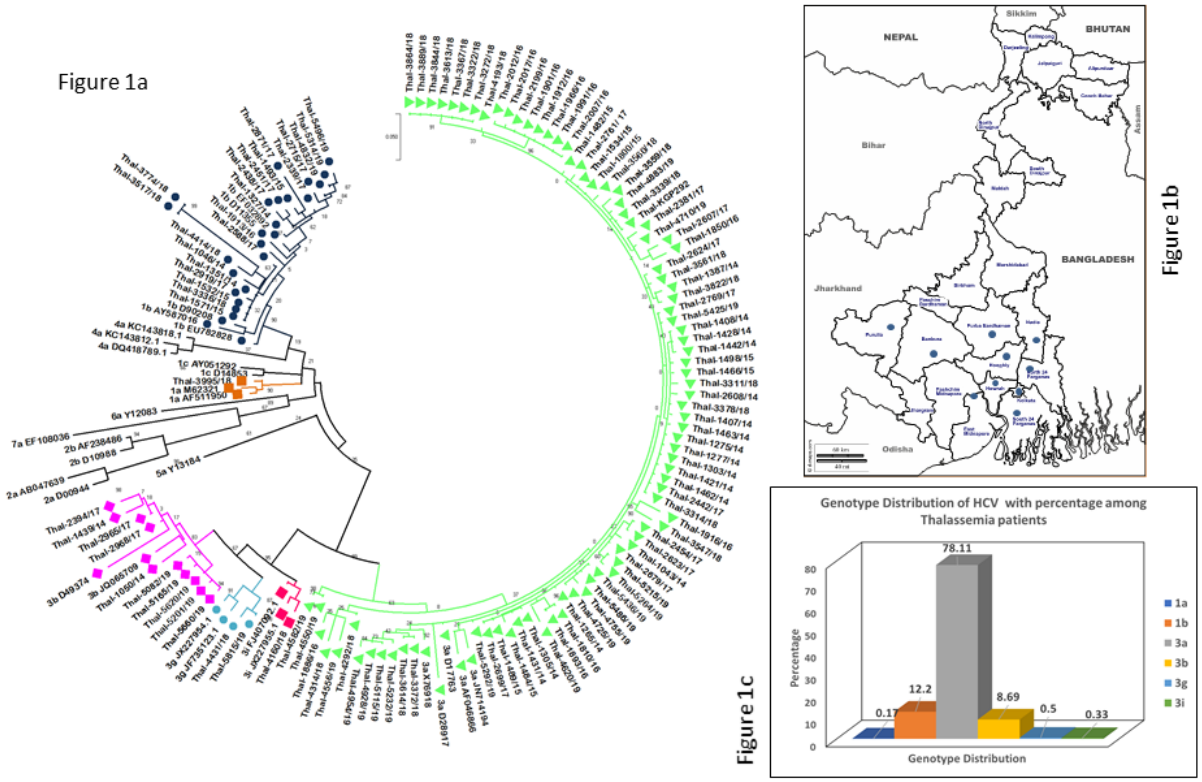

Figure 1: Phylogenetic tree (Fig 1a), Relative positions of Transfusion Centres in West Bengal Map (Fig 1b), Genotype distribution of HCV with percentage among thalassemia patients (Fig 1c)

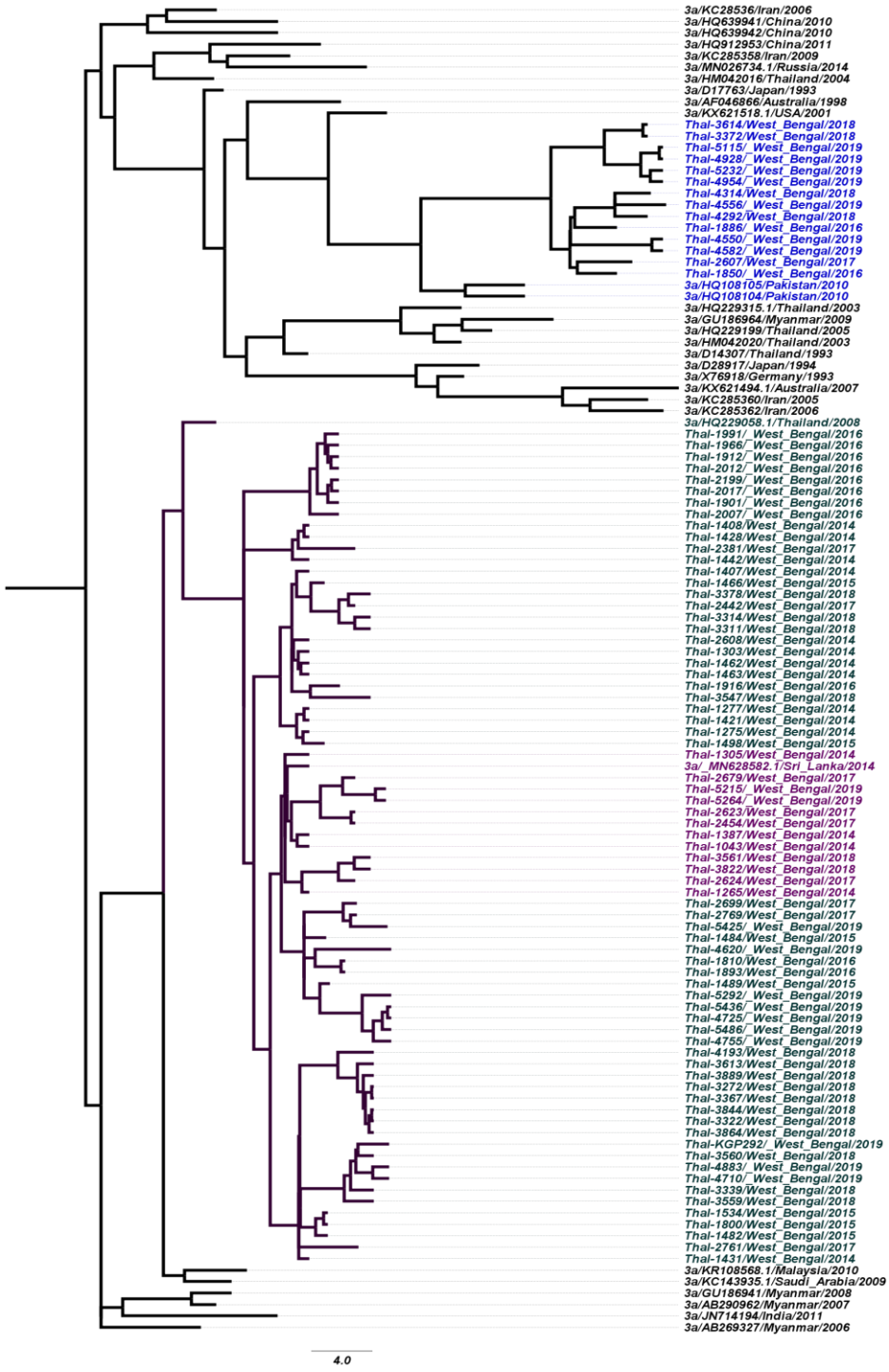

Figure 2: 3a phylogeography

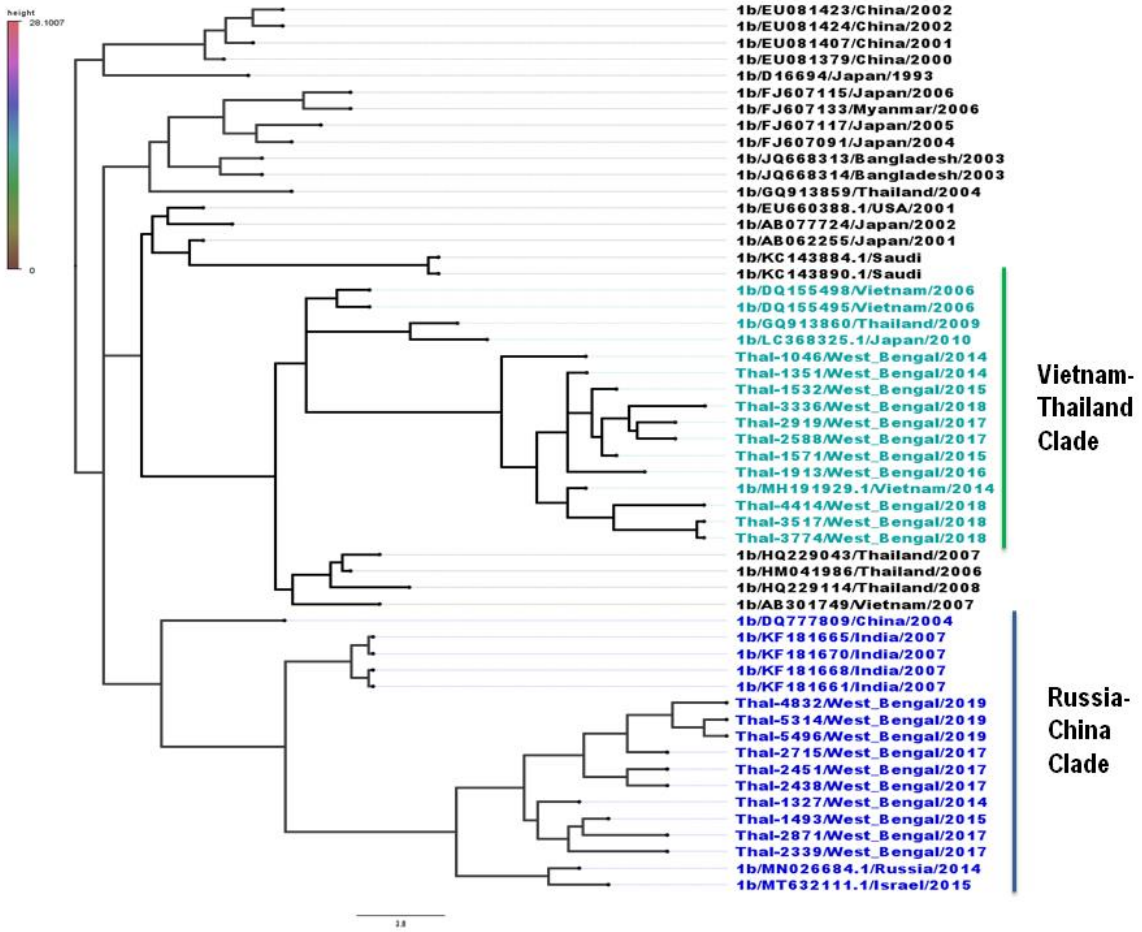

Figure 3: 1b phylogeography