

## RESEARCH ARTICLE

# Hepatitis C virus genotype/subtype distribution and evolution among Chinese blood donors: Revealing recent viral expansion

Yu Zhang<sup>1,2</sup>, Zhan Gao<sup>2,3</sup>, Shaoli Wang<sup>2,3</sup>, Jing Liu<sup>4</sup>, Ness Paul<sup>4</sup>, Tao He<sup>5</sup>, Cunxu Liu<sup>6</sup>, Hongbin Zhang<sup>7</sup>, Yunlai Lv<sup>8</sup>, Ru'an Cao<sup>9</sup>, Wei Mao<sup>5</sup>, Jianhua Wan<sup>7</sup>, Hongli Ma<sup>8</sup>, Mei Huang<sup>9</sup>, Yu Liu<sup>2,3</sup>, Jingxing Wang<sup>2,3</sup>, Pu Liao<sup>10</sup>, Peibin Zeng<sup>11\*</sup>, Miao He<sup>2,3\*</sup>, Hua Shan<sup>12</sup>

**1** The Fourth Affiliated Hospital Zhejiang University School of Medicine, Yiwu, China, **2** Institute of Blood Transfusion, Chinese Academy of Medical Sciences, Chengdu, China, **3** Sichuan Blood Safety and Blood Substitute International Science and Technology Cooperation Base, Chengdu, China, **4** The Johns Hopkins Medical Institutions, Baltimore, MD, United States of America, **5** Chongqing Blood Center, Chongqing, China, **6** Guangxi Blood Center, Liuzhou, Guangxi, China, **7** Urumqi Blood Center, Urumqi, Xinjiang, China, **8** Luoyang Blood Center, Luoyang, Henan, China, **9** Mianyang Blood Center, Mianyang, Sichuan, China, **10** The People's Hospital of Chongqing, Chongqing, China, **11** West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan, China, **12** Stanford University, Stanford, CA, United States of America

\* These authors contributed equally to this work.

\* [zengpeibin@linve.cn](mailto:zengpeibin@linve.cn) (PZ); [hemiao@ibt.pumc.edu.cn](mailto:hemiao@ibt.pumc.edu.cn) (MH)



## OPEN ACCESS

**Citation:** Zhang Y, Gao Z, Wang S, Liu J, Paul N, He T, et al. (2020) Hepatitis C virus genotype/subtype distribution and evolution among Chinese blood donors: Revealing recent viral expansion. *PLoS ONE* 15(7): e0235612. <https://doi.org/10.1371/journal.pone.0235612>

**Editor:** Jason Blackard, University of Cincinnati College of Medicine, UNITED STATES

**Received:** December 28, 2019

**Accepted:** June 19, 2020

**Published:** July 10, 2020

**Copyright:** © 2020 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** The study was supported by the Recipient Epidemiology and Donor Evaluation Study-III (REDS-III) International Component (China) (HHSN2682011000081) and the CAMS Innovation Fund for Medical Sciences (CIFMS) (CAMS-2016-I2M-3-025).

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Hepatitis C virus (HCV) genotype (GT) distribution in China shows significant geographical and demographic difference. As a routinely tested virus in Chinese blood bank systems, rare molecular epidemiology research in blood donors is reported. Our purpose is to investigate the HCV GT/subtypes distribution, phylogenetic analysis and population genetics in Chinese blood donors. Anti-HCV screen positive samples and donor demographics were collected. HCV Core and E1 gene fragments were amplified by RT-PCR, followed by sequencing and phylogenetic analysis to determine HCV GTs/subtypes using MEGA 7.0. The population genetics were performed using Arlequin v3.0 and Beast v1.10.4. SPSS Statistics 17.0 software was used to analyze the correlation between HCV GTs/subtypes distribution and demographic characteristics. 419 and 293 samples based on Core and E1 gene respectively were successfully amplified. HCV 1a, 1b, 2a, 3a, 3b, 6a, 6e and 6n were found, and the corresponding proportions were 0.66% (3/455), 58.68% (267/455), 17.80% (81/455) and 5.05% (23/455), 3.52% (16/455), 12.31% (56/455), 0.88% (4/455) and 0.66% (3/455). Samples from Guangxi showed the most abundant genetic diversity with 8 subtypes were found. The number of haplotypes in HCV-1b is higher than 2a and 6a. The negative Tajima's *D* and Fu's *F<sub>s</sub>* values of HCV-1b, 2a and 6a suggested the population expansion of those HCV subtypes. The distribution of HCV GT showed significant statistical difference by age and ethnicity. **Conclusion:** An abundance of HCV genetic diversity was found in Chinese blood donors with mainly 1b and then 2a subtype. There were significant geographical and demographic differences in HCV GTs/subtypes among Chinese blood donors. HCV

subtype 1b has stronger viability and HCV subtype 6a has experienced significant expansion.

## Background

Hepatitis C Virus (HCV) is a single positive strand RNA virus with envelope belonging to the Flaviviridae family. It can cause both acute and chronic hepatitis, ranging in severity from a mild illness lasting a few weeks to a serious, lifelong illness including Liver cirrhosis, liver cancer and liver failure [1]. WHO has estimated 71 million people have chronic hepatitis C infection. In China, the proportion of new HCV cases is 0.06 ‰ per year, which is even higher in some regions such as Fujian, reaching 6.01% in 2010 [2].

Hepatitis C is mainly transmitted via blood transfusion and plasma derivatives usage before Blood Donation Law initiated in 1998. With the Chinese government banned the paid blood donation, the prevalence of HCV remains stable after a significant reduction in blood donors. Nowadays, HCV infections caused by sexual transmission has been increasing continuously in recent years. [3,4]. While the initiation of HCV infection is concealed, it still remains as a huge threats to Chinese blood safety.

With high variable genome, the diversity of HCV nucleic acid sequence is as high as 30%, and according to which, HCV is divided into 7 GTs and more than 90 subtypes [5]. HCV GT distribution shows geographical differences: HCV-1, HCV-2 and HCV-3 are widely distributed globally; The major HCV subtypes include 1b, 2a, 3a, 3b and 6a in China, of which subtype 1b accounts for more than 70%, followed by 2a [5]; 1b and 2a are more common in north and west China, while 3a, 3b and 6a are more popular in southwest and south China. Besides, recently, some novel group of HCV variants has been found in China [6].

The distribution of HCV GT is diverse in different subgroups such as blood donors, intravenous drug users (IDU) and dialysis patients. In the group of blood donors and blood recipients, the popular subtypes were HCV-1b and 2a [7] but had shown a decrease in some regions with HCV-3 and 6 on the rise [7]. HCV-2a was also found in people who have high-risk sexual behaviors or unsafe medical injection [8].

Characteristics of HCV-infected people among Chinese blood donors were reported [9–11]. In recent years, with the increasing population mobility and the diversified transmission modes, the distribution of HCV GT in blood donors of China is undergoing significant changes [12]. An update on HCV GT distribution among blood donors is needed to re-evaluate the blood screening strategy in blood screening to avoid unnecessary cost and waste of collected units due to false positive results [13]. In this study, plasma samples from anti-HCV positive blood donations were collected from five Chinese blood centers (Luoyang, Chongqing, Mianyang, Urumqi and Guangxi) in three consecutive years (2012–2014). Our aims were: 1) To understand HCV GT and subtype distribution in Chinese blood donors; 2) to investigate the expanding trends of HCV GTs/subtypes using HCV phylogenetic and population dynamics analysis; 3) To provide molecular epidemiology supporting data to improve nucleotide acid test (NAT) assay development based on HCV genome diversity, and to inform public health policy making regarding HCV prevention and control.

## Materials and methods

All methods were carried out in accordance with relevant guidelines and regulations, and The Ethical Review Committee (IORG0006536) of Chinese Academy of Medical Sciences/Peking

Union Medical College approved the study and the ethics approval number was X101222002. Before sampling, the voluntary blood donors have signed an informed consent form.

### Study population

802 anti-HCV positive plasma samples were collected from 5 Chinese blood centers including: Chongqing Blood Center, Chongqing; Urumqi Blood Center, Urumqi, Xinjiang; Luoyang Blood Center, Luoyang, Henan; Mianyang Red Cross Blood Center, Mianyang, Sichuan; Guangxi Blood Center, Liuzhou, Guangxi. Samples were collected with EDTA anticoagulation tube, stored at  $-20^{\circ}\text{C}$ , and then transported to the Institute of Blood Transfusion, Chinese Academy of Medical Sciences through the cold chain.

The study population was divided into different age groups: 18–20, 21–30, 31–40, 41–50, >50 years of age; Ethnicity group: Han, Hui, Uygur, Zhuang; Occupation groups: student, commercial, worker, peasant, working at home, other; Educational levels:  $\leq$ primary school, middle or high school, >high school, other, missing.

### HCV RNA extraction and reverse-transcription polymerase chain reaction (RT-PCR)

The required 8 primers including outer and inner amplification were designed after reference [14]. At present, three most commonly used HCV coding regions to genotyping are Core, E1 and NS5B [15–17], which represents high accuracy and sensitivity, however, sometimes different HCV subtypes can be obtained with the above three regions for the same HCV genome, suggesting the presence of mixed infection or gene recombination [18,19]. In our study, coding regions of Core and E1 were separately used to reconstruct independent HCV GTs/subtypes phylogenetic trees in order to make more accurate genotyping, and the final amplified HCV products based on Core and E1 gene fragments were 401bp and 592bp, respectively.

HCV RNA was extracted from 140 $\mu\text{L}$  of each plasma sample by the principle of Membrane adsorption using the QIAmp Viral RNA Mini kit (QIAGEN, German) following the instructions strictly, and then it was reverse transcribed into DNA fragment with the help of antisense outer primers of Core and E1 using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Switzerland). The nested PCR based on Core and E1 gene fragment was performed using 5 $\mu\text{L}$  of the RT-PCR product with 2 $\times$ TSINGKE MASTER MIX (TSINGKE Biotech Co Ltd, Chengdu, China), and primers as specified previously [20]. The final PCR products were confirmed by 2% agarose gel electrophoresis and visualization with ethidium bromide staining, and amplified products were sent to TSINGKE Biotech Co Ltd (Chengdu, China) for purifying and sequencing with the sense and antisense inner primers.

### HCV genotyping, phylogenetic and evolutionary analysis

All the sequences of the Core and E1 regions were checked and spliced using Mega 7.0, and the spliced HCV sequences were aligned with HCV reference sequences (S1 Table). Phylogenetic analysis was performed using the neighbor-joining and the Kimura two-parameter method to get the accurate HCV genotyping. Node supports were evaluated with 1000 bootstrap replicates. HCV GT/subtype was identified when it was included in a monophyletic cluster with classified reference strains.

Arlequin v3.0 was used to perform population genetics data analysis including calculating the haplotype frequencies, estimating the past population expansion of species by Tajima's  $D$  and Fu's  $F_s$  value, and if the values are negative and  $P < 0.05$  or  $0.01$ , representing the population is considered to have undergone expansion. BEAST v1.10.4 was used to get the Bayesian skyline plot (BSP) to analyze growth history of different HCV subtypes. Uniform prior was

selected and the evolutionary rates of Core and E1 fragment for each dominant subtypes were shown in S2 Table.

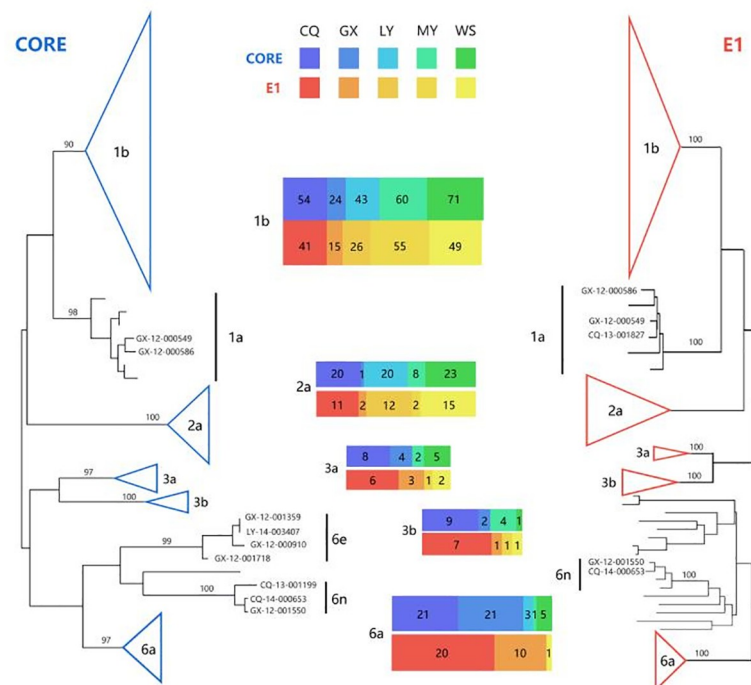
## Statistical analysis

Statistical data analyses were performed with SPSS Statistics 17.0. The Pearson Chi-Square test was used to compare the ratio of risk factors between HCV GT/subtypes and demographic characteristics (age, gender, ethnic, occupation, educational level), and P value < 0.05 was considered to be statistically significant.

## Results

### HCV genotyping and phylogenetic analysis

Among the 802 anti-HCV positive plasma samples, 419 and 293 were successfully amplified based on Core gene fragment and E1 gene fragment correspondingly. The determined genotype among samples that were successfully amplified by both Core and E1 fragment were consistent, and 36 subtypes, that were successfully amplified by E1 fragment, were not amplified by Core fragment. Based on the above, 455 of 802 samples based on Core and (or) E1 fragment were successfully amplified. The successfully amplified samples were then sequenced and used to construct phylogenetic trees (Fig 1). In detail, the predominant HCV subtype was 1b (58.68%, 267/455), followed by subtype 2a (17.80%, 81/455). Other HCV subtypes including 1a, 3a, 3b, 6a, 6e and 6n were also found among these blood donors, and the proportion of the



**Fig 1. Phylogenetic tree constructed using HCV Core and E1 gene sequences from blood donors in China.**

Phylogenetic analysis was performed using the neighbor-joining and the Kimura two-parameter method to get the accurate HCV genotyping, and the node supports were evaluated with 1000 bootstrap replicates. The colored triangles and vertical lines represent different HCV subtypes. The upper bar refers to the Core region and the lower bar refers to the E1 region. The colored rectangular area represents the number of HCV subtypes in five regions including Chongqing (CQ), Guangxi (GX), Luoyang (LY), Mianyang (MY) and Urumqi (WS).

<https://doi.org/10.1371/journal.pone.0235612.g001>

above subtypes were 0.66% (3/455), 5.05% (23/455), 3.52% (16/455), 12.31% (56/455), 0.88% (4/455) and 0.66% (3/455), respectively.

The distribution of HCV GT among blood donors in China showed significant geographical difference ( $P < 0.05$ ) (S3 Table). The predominant HCV subtypes in different regions were as followed: 1b (59.72%, 43/72) and 2a (34.72%, 25/72) in Luoyang, 1b (46.21%, 61/132) and 6a (18.94%, 25/132) in Chongqing, 1b (80.77%, 63/78) and 2a (10.26%, 8/78) in Mianyang, 1b (67.57%, 75/111) and 2a (22.52%, 25/111) in Urumqi, and 1b (41.94, 26/62) and 6a (35.48%, 22/62) in Guangxi.

## HCV GT and demographic characteristics

Table 1 showed the association between some HCV subtypes (1b, 2a, 3a, 3b and 6a) and gender, age, ethnicity, education status and occupation. There were significant differences between HCV GT and age and ethnicity ( $P < 0.05$ ), while no statistical differences were observed between HCV GT and gender, occupation and education level ( $P$  value  $> 0.05$ ). In the age groups of 21–30, 41–50 and  $>50$ , HCV subtypes 1b and 2a accounted for a higher proportion, while in the age groups of  $\leq 20$  and 31–40, 1b and 6a were more common. In Han population, HCV 1b and 2a were the prevalent subtypes, accounting for 72.6% (312/430) of the whole ethnic groups, while HCV 6a was more common in Zhuang population.

## HCV population dynamics

The dominant subtypes were 1b, 2a, and 6a. The subtypes determined by Core gene were 251, 72 and 51, respectively and the corresponding haplotypes of Core 1b, 2a and 6a were 178, 55 and 44; The sequences of 1b, 2a, and 6a determined by E1 gene were 186, 42 and 31, respectively and the corresponding haplotypes of E1 1b, 2a and 6a were 183, 42 and 31 (Table 2). The neutral test values of Tajima's  $D$  and Fu's  $F_s$  of each HCV subtypes were negative, the  $P$  value corresponding to Fu's  $F_s$  of each subtype was  $< 0.01$ , and the  $P$  value corresponding to Tajima's  $D$  of subtype 1b was  $< 0.05$ ; The Tajima's  $D$  and Fu's  $F_s$  values of HCV 6a were both negative and  $P$  value is less than 0.01 (Table 2).

The spatio-temporal changes estimated in the HCV population size were illustrated as results of BSP shown in Fig 2A and 2B. For HCV-1b the Core and E1 genes showed mutually rapid population growth between 1992 and 1998. While the Core gene for HCV-1b showed another period of continuous population growth until 2005. As for HCV-2a, both Core and E1, the genes mutually showed population growth after 2000. HCV-6a was estimated as a new member of the blood donor population that appeared around 1995 and the sudden population expansion occurred between 2002 and 2010, based on the calculation of BSP of the Core and E1 genes.

## Discussion

### Regional differences

In this study, HCV GT/subtype distribution in groups of blood donors showed significant geographical variations that are different from findings in previous publications.

In Guangxi, two studies performed in 1990s showed that the prevalent GTs were 1b and 1a, 3b and 1b [21,22], respectively. However, our study together with reports in recent years, suggest that HCV 1b, 6a has become the dominant subtype over time in Guangxi, which is consistent with a meta-analysis reported [12]. In addition, we also found that the proportion of HCV-6a increased compared with other previous reports in Guangxi while the proportion of HCV-2a were decreasing [21–26]. In early years of 1990s, HCV-2b has been detected in

**Table 1. The association between HCV genotypes/subtypes and demographic characteristics.**

HCV		GTs	HCV-1b (n = 267)	HCV-2a (n = 81)	HCV-3a (n = 23)	HCV-3b (n = 16)	HCV-6a (n = 56)	P*
Demographic characteristics								
Gender	Male/female		162/105	46/35	13/10	10/6	29/27	>0.05
Age	≤20		3 (75)	0	0	0	1 (25)	<0.05
	21–30		89 (63.12)	32 (22.70)	9	3	9	
	31–40		36 (46.15)	4	8	6	21 (26.92)	
	41–50		106 (61.99)	26 (15.20)	4	6	23	
	>50		33 (57.89)	19 (33.33)	2	1	2	
Ethnicity	Han		239 (60.97)	73 (18.62)	19	13	42	<0.05
	Hui		3 (60.00)	1	0	1	0	
	Uygur		14 (70.00)	3	3	0	0	
	Zhuang		3 (14.29)	1	1	1	13 (61.90)	
Education status	>High school		74	26	7	2	14	>0.05
	Middle or high school		159	45	11	14	29	
	≤Primary school		4	3	0	0	4	
Occupation	Student		36	10	2	0	4	>0.05
	Commercial		36	9	3	2	6	
	Worker		24	13	2	2	8	
	Peasant		51	14	2	1	9	
	Working at home		64	21	7	6	15	

\* using method of Fisher's exact test

The values in brackets mean the higher proportion of each group.

<https://doi.org/10.1371/journal.pone.0235612.t001>

Guangxi [21,22], and since then, this subtype hasn't never been found in later studies including this one. HCV-6n was detected which was never been reported before in Guangxi. Probably due to the special geographical environment of Guangxi and its proximity to Hong Kong, Macao, Vietnam, Thailand and other countries and regions, such as the HCV transmission of rare genotypes in these countries and regions; coupled with the wide communication leading to HCV genotype in Guangxi extremely rich and complicated.

In Mianyang (Sichuan province), Luoyang (Henan province) and Urumqi (Xinjiang province), the prevalent genotypes were both HCV 1b and 2a. According to the previous reports, the major two HCV subtypes in Sichuan were 1b and 3b or 3a [27–30]. In our study, however, HCV-1b and 2a became the predominant subtypes in Mianyang with the prevalence rate of 1b

**Table 2. HCV population genetic characteristics in China.**

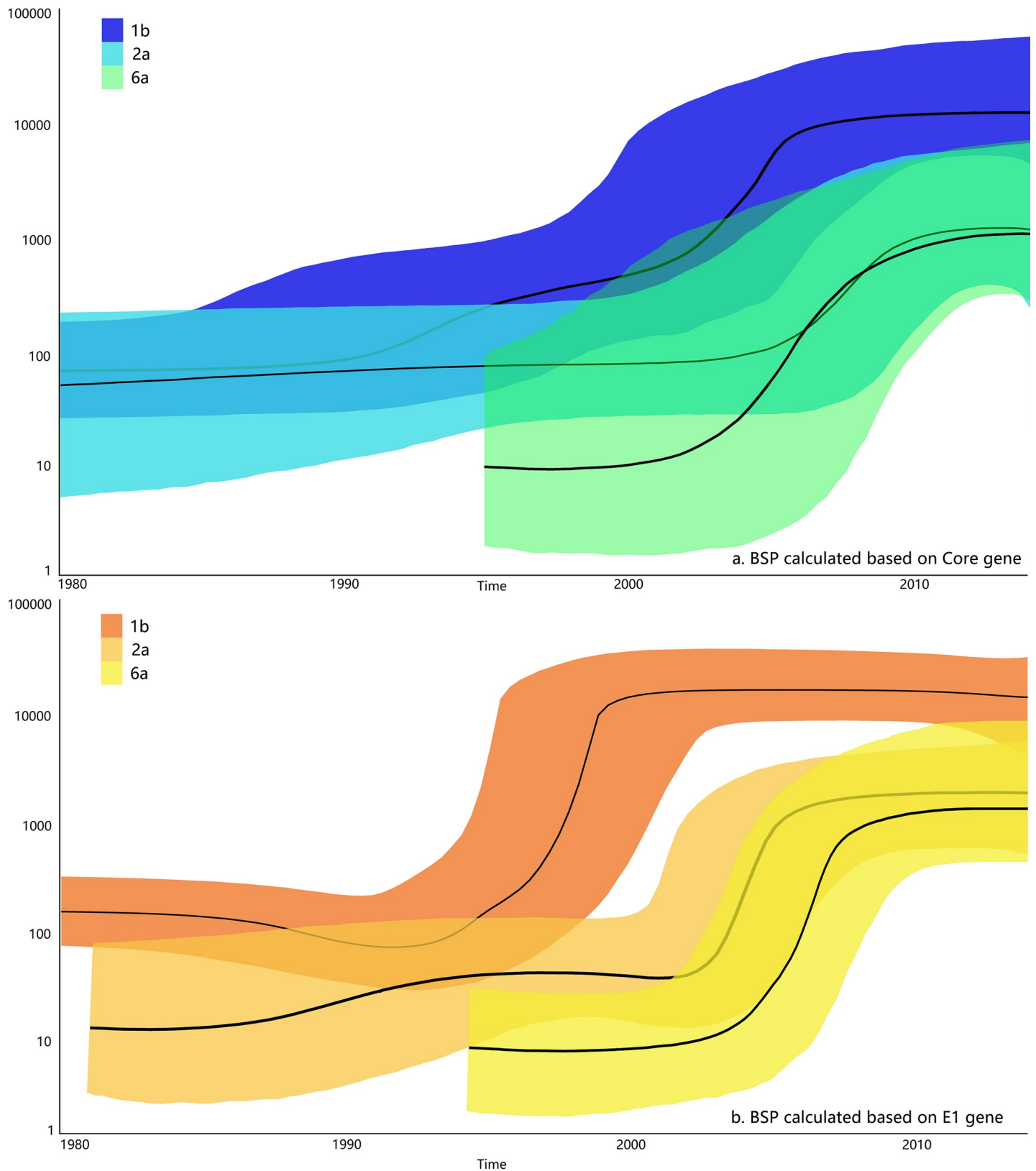
HCV GTs	n	Haplotypes	Tajima's D	P	Fu's Fs	P
Core-1b	251	178	-1.48031	0.025*	-24.06017	0.002**
Core-2a	72	55	-0.83021	0.204	-24.20050	0.000**
Core-6a	51	44	-1.80712	0.009**	-25.44373	0.000**
E1-1b	186	183	-0.87571	0.213	-23.75550	0.005**
E1-2a	42	42	-0.10326	0.549	-14.12353	0.001**
E1-6a	31	31	-0.88458	0.186	-11.69440	0.003**

\*P < 0.05 significant population expansion

\*\*P < 0.01 highly significant population expansion.

<https://doi.org/10.1371/journal.pone.0235612.t002>





**Fig 2. Population genetic characteristics of HCV with different genotypes among blood donors in China.** The solid black part is the median of the estimated changes in genetic diversity. Other colored parts is the 95% confidence interval. The horizontal axis is the year. The vertical axis is  $N_{er}$  (the effective population size of

the experimental sample). a: The BSP for HCV 1b(blue), 2a(light blue) and 6a(green) based Core gene. b: The BSP for HCV 1b(orange), 2a(light orange) and 6a(yellow) based E1 gene.

<https://doi.org/10.1371/journal.pone.0235612.g002>

up to 80%, and the prevalence of 2a was significantly higher than 3a and 3b. In Urumqi, the prevalence of two major subtypes was consistent with three previous reports [31–33]. With limited previous studies about HCV genotypes, the changing trend of each HCV GT in the above two regions can't be easily found over time. In Luoyang, the prevalent HCV subtypes were consistent with most of previous studies reported, while HCV-2a exhibits a trend of slight increase.

In Chongqing, the distribution of predominant HCV subtypes showed high gene abundance that HCV-1b and 2a [34–36], 1b and 3b [37,38], 2a and 3a [39] have ever been reported as the major HCV subtypes. Comparing with a 2003–2004 study performed in blood donors of Chongqing [39], the proportion of HCV 6a has increased significantly, which was consistent with pooled results of a meta-analysis [12].

The distribution of HCV subtypes had undergone some changes over time in some places of China, just as what we have found in Guangxi, Chongqing and Mianyang. The emergence of new HCV subtypes and changing trends of HCV GTs may be ascribed to: 1) Increasing population mobility: With the increase of domestic or international trade and communication, the interaction between people will naturally increase, which may lead to the spread of HCV; 2) The changing transmission routes: In recent years, there was an increase in HCV IDUs, possibly due to common syringes contaminated by blood, which may represent an increasing of HCV subtypes 3 and 6 [8]. In our study, the increasing proportion of HCV-6a in Guangxi and Chongqing may be a consequence of IDU related spreading in these two areas [25,38]; 3) The improved sensitivity and specificity of testing methods and the sample size. Most notably, a multi-center large sample research still may need to be performed to reveal more meaningful results showing HCV changing trends.

## Demographic association

Reports about the association between HCV GTs and host demographics in China were inconsistent [8,40]. Subtype 1b was higher in male than in female in some reports [41]. The possible reason may be that gender difference exhibits different immune response against HCV GTs and the life style of male population is complex and thus the chances of infecting HCV are relatively high. Previous studies found that in the older group ( $\geq 40$ ), HCV-1b and 2a were the predominant subtypes, while in the younger group, the proportion of HCV-6 and 3 showed an increasing trend with drug use or high-risk sexual behavior by young people [42,43].

In our study, significant statistical result was found between HCV GTs and age. HCV subtypes 1b and 2a are more common in the age groups of 21–30 and  $>41$ , while 1b and 6a are more common in the younger groups of 18–20 and 31–40, which was consistent with early reports, while higher proportion of 1b and 2a was found for the group of 21–30. We speculate that it may take years from most HCV infected individuals to develop clinical symptoms [44], thus HCV 1b and 2a still remained the main subtypes, and in addition, sampling deviation should also be taken into account.

In Han population, 1b and 2a accounted for 72.6% of the whole ethnic groups, while HCV 6a is more common in Zhuang population, and the possible reasons are: 1) The lifestyle of different nationalities are diverse, and HCV transmission modes are also diverse [45]; 2) different ethnicities show different immune responses to HCV infection with different subtypes [46]. In addition, there was no significant association between HCV GTs and occupation, educational level in our study, which was similar to the previous reports [11,47]. The sample size in this



study is large and representative; however, the exact mechanism of the relationship between HCV GT distribution and demographic characteristics needs more investigation.

### Population dynamics

BSP analyses were performed using the sequences determined in this study and it showed that HCV 1b, 2a and 6a have experienced different rates of population expansion at different times in the past.

**BSP for HCV 1b.** As shown in BSP 1b based on Core and E1 gene fragment, that they concurrently highlighted the growth in the HCV-infected population size from 1992 to 1998, consistent with a period when an official plasma donation campaign and inadequate donor screening protocol resulted in contaminated blood supply and infected blood donors. This 1990s campaign in Henan province resulted in more than 500,000 blood donors infected with HCV [48–51]. To eliminate this risk, the central government of China banned and outlawed paid blood donation in 1998 [49], and this corresponding change was well reflected in our Core and E1 1b BSP that since 1998, this population expansion tended to be stable. In addition, we found that Core-1b population still through transient expansion after 2000 (2000–2005). We do not deny different sample sizes, evolution rate estimation, parametric model and other aspects resulting this divergence, but a more reasonable explanation, we think, may be that although the enactment of the blood donation law and corresponding knowledge publicity of blood donation reduced the HCV spreading to a certain extent, after all, the infected population is large, and together with population mobility and other factors, allowing wide HCV dissemination. However, in this study, based on BSP results, we believe that HCV 1b is the specific HCV subtype who was the prime culprit of the huge disaster to Chinese blood safety. After the sudden expansion of 1b population, HCV 1b has remains as the main HCV subtype distributed in China.

**BSP for HCV 2a.** Previous reports have showed that unsafe medical practices have a certain relationship with HCV-2a [52]. Since 2004, HCV outbreaks in Anhui, Hubei, and Fujian due to the reuse of unsterilized or inappropriately sterilized needles have been reported [53,54], consistent with the rapid expansion of HCV 2a population in 2005 with BSP. Our BSP based on Core and E1 gene showed that HCV-2a population expanded from 1990 to 2010 at different rates, and according to the above cases of unsafe medical practices, we speculate that HCV transmission occurred from time to time with improper medical behaviors especially in some poor or less-developed regions, resulting in the expansion of this population. In addition, since 2a is the prevalent HCV GT in HCV infection transmitted by blood transfusion [11], and HCV infection caused by blood transfusion or blood products may cause expansion of HCV-2a population [55]. These findings emphasize the importance of tracking HCV outbreaks caused by improper needle reuse or unsterilized injections in order to reduce or eliminate further HCV infection.

**BSP for HCV 6a.** Previous studies have reported that HCV-6 was more frequent among IDU patients [56–58]. Before 2000, transfusion of contaminated blood or blood products was the main route of HCV transmission in China, however, some risk factors such as using illegal intravenous drugs, unsafe medical practices and risky behaviors of homosexual men have become the main HCV transmission routes now [59], which greatly promoted the spread of HCV and also built the best conditions for rapid HCV molecular evolution [60,61]. Our study showed that HCV-6a population undergone population expansion in 2000–2010, which is also corresponding with the above reports [24–26] that the proportion of HCV-6a in Chongqing and Guangxi increased. This may represent that HCV-6a had a tendency of expanding to the general population including blood donor. In addition, HCV-6a expanded relatively more recently compared with the HCV-1b populations.

This study has its limitation in that samples successfully amplified based on the E1 were less than Core, resulting from degradation of nucleic acid with E1 gene fragments being amplified at a later time after Core and this experiment was carried out among blood donors.

In summary, the haplotype diversity in the population is an indicator to measure the population genetic diversity, and the more haplotype the population has, the richer genetic diversity is and the stronger the adaptability of the species to the environment will be. In this study, the number of haplotypes in each HCV subtypes is large, and subtype 1b is higher than 2a and 6a from the two gene fragments, which indicates that HCV subtype 1b has stronger viability. HCV population genetic characteristics showed HCV-1b, 2a and 6a all experienced a certain degree of population expansion, which is consistent with BSP results, and notably, HCV-6a had a significant expansion in China. Therefore, we suggested that relevant health departments should take more effective measures to control HCV infection, especially strengthening supervision of IDU and publicizing self-protection of risky behaviors of homosexual men, and in addition, it is necessary to organize large-scale molecular epidemiological investigation and to give early warning of HCV outbreak timely.

China Center for Disease Control and Prevention (CCDC) reports that the prevalence of hepatitis C in China has fallen to around 0.4% after the shift to more safe medical practices and illegal commercial blood donations[62]. Therefore, keeping strengthening the management of blood donors and screening of donations is a good way to eliminate HCV-infected people to donate. However, new HCV subtypes or variants will emerge just as HCV-6a has been found in donor group, which may challenge the eligibility of screening strategy and assay. So, continuous surveillance of HCV GT/subtype and population dynamics changing are important to control the risk of transfusion-transmitted HCV in China.

## Supporting information

**S1 Table. Reference sequences of HCV subtypes used in phylogenetic analysis.**

(DOCX)

**S2 Table. Evolutionary rates of HCV subtypes used for Bayesian skyline plot analysis.**

(DOCX)

**S3 Table. The geographical distribution of HCV subtypes among blood donors in China.**

(DOC)

## Acknowledgments

The study was supported by the Recipient Epidemiology and Donor Evaluation Study-III (REDS-III) International Component (China) (HHSN268201100008I) and the CAMS Innovation Fund for Medical Sciences (CIFMS) (CAMS-2016-I2M-3-025).

## Author Contributions

**Data curation:** Peibin Zeng, Miao He.

**Formal analysis:** Yu Liu.

**Funding acquisition:** Jingxing Wang.

**Methodology:** Peibin Zeng, Miao He.

**Project administration:** Yu Zhang, Shaoli Wang.

**Resources:** Tao He, Cunxu Liu, Hongbin Zhang, Yunlai Lv, Ru'an Cao, Wei Mao, Jianhua Wan, Hongli Ma, Mei Huang.

**Software:** Pu Liao.

**Supervision:** Zhan Gao, Miao He.

**Validation:** Ness Paul, Miao He, Hua Shan.

**Visualization:** Jing Liu, Jingxing Wang, Hua Shan.

**Writing – original draft:** Yu Zhang.

## References

1. Zhang LX, An Y, Zhang XG (2011) The HCV genotype distribution among chronic HCV infectors in Shandong. *Chinese journal of pathogen biology* 6: 567–570.
2. Bennett H, Waser N, Johnston K, Kao JH, Lim YS, et al. (2015) A review of the burden of hepatitis C virus infection in China, Japan, South Korea and Taiwan. *Hepatology* 9: 378–390. <https://doi.org/10.1007/s12072-015-9629-x> PMID: 26071238
3. Wang ZY, Li Y, Xu S, Li GW, et al. (2019) Genotype Distribution of Hepatitis C in Chuxiong, Yunnan. *China Modern Doctor* 57: 5–11+15+169.
4. Daniels D, Grytdal S, Wasley A, Centers for Disease C, Prevention (2009) Surveillance for acute viral hepatitis—United States, 2007. *MMWR Surveill Summ* 58: 1–27.
5. Ye Y, Yan YS, Chen G, Yan PP, Zheng WX, et al. (2013) Molecular epidemiology of hepatitis C virus among different groups of people in the province of Fujian, China. *Arch Virol* 158: 611–618. <https://doi.org/10.1007/s00705-012-1513-6> PMID: 23132411
6. An Y, Wu T, Wang M, Lu L, Li C, et al. (2014) Conservation in China of a novel group of HCV variants dating to six centuries ago. *Virology* 464–465: 21.
7. Wu LP, Liu YP (2001) Research on HCV genotype among different group people in Shanxi. *The Chinese journal of microbiology and immunology* 21: 14–17.
8. Peng J, Lu Y, Liu W, Zhu Y, Yan X, et al. (2015) Genotype Distribution and Molecular Epidemiology of Hepatitis C Virus in Hubei, Central China. *PLoS One* 10: e0137059. <https://doi.org/10.1371/journal.pone.0137059> PMID: 26325070
9. Lin R, L YF, Chen WR, Wen JS, Wang B (2014) Hepatitis C virus infection and genotypes among voluntary blood donors in Wenzhou. *China Modern Doctor* 52: 151–153.
10. Gao XF, Cui QA, Shi X, Su J, Peng ZH, et al. (2011) Prevalence and trend of hepatitis C virus infection among blood donors in Chinese mainland: a systematic review and meta-analysis. *Bmc Infectious Diseases* 11.
11. Xu JF, Zhu SW, Jiang NZ (2016) Investigation on HCV epidemiology among voluntary blood donors in Jiangsu. *Journal of Nanjing Normal University (Natural Science Edition)*: 73–77.
12. Zhang Y, Chen LM, He M (2017) Hepatitis C Virus in mainland China with an emphasis on genotype and subtype distribution. *Virology* 14: 41. <https://doi.org/10.1186/s12985-017-0710-z> PMID: 28231805
13. Fu P, Lv Y, Zhang H, Liu C, Wen X, et al. (2019) Hepatitis C virus prevalence and incidence estimates among Chinese blood donors. *Transfusion* 59: 2913–2921. <https://doi.org/10.1111/trf.15432> PMID: 31271469
14. Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, et al. (2005) Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* 42: 962–973. <https://doi.org/10.1002/hep.20819> PMID: 16149085
15. Bukh J, Miller RH, Purcell RH (1995) Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis* 15: 41–63. <https://doi.org/10.1055/s-2007-1007262> PMID: 7597443
16. Zhang Z, Yao Y, Wu W, Feng R, Wu Z, et al. (2013) Hepatitis C virus genotype diversity among intravenous drug users in Yunnan Province, Southwestern China. *PLoS One* 8: e82598. <https://doi.org/10.1371/journal.pone.0082598> PMID: 24358211
17. Tagnoukam-Ngoupo PA, Ngoufack MN, Kenmoe S, Lissack SF, Amougou-Atsama M, et al. (2019) Hepatitis C virus genotyping based on Core and NS5B regions in Cameroonian patients. *Virology* 16: 101. <https://doi.org/10.1186/s12985-019-1214-9> PMID: 31399103
18. Xu R, Wang M, Liao Q, K H (2015) Confirmation of a Mixed Infection with Different Genotypes/Subtypes of HCV. *Modern Medicine* 8.

19. Zhang KY, Zhou LY, Yang XJ, HY Z (2014) Study on gene recombination of hepatitis C virus in Asia. *Journal of Lanzhou University*: 13–17.
20. Lu L, Nakano T, He Y, Fu Y, Hagedorn CH, et al. (2005) Hepatitis C virus genotype distribution in China: predominance of closely related subtype 1b isolates and existence of new genotype 6 variants. *J Med Virol* 75: 538–549. <https://doi.org/10.1002/jmv.20307> PMID: 15714489
21. Ge XM, Wang SS (1996) The HCV serological genotype and molecular biology research of the genotype among IDUs, HCV infectors and blood donors in Guangxi. *Chinese Journal of Experimental and Clinical Virology* 10: 312–317.
22. Ge XM, Wang SS (1996) HCV infection among IDUs in Guangxi and its molecular biology research. *Guangxi Journal of Preventive Medicine* 2: 1–7.
23. Yin Q, Yang QL, Wang XY (2016) Analysis of HCV genotype and genetic affinity in Liuzhou, Guangxi. *china tropical medicine* 16: 965–969.
24. Zhu MY, Su MH, Jiang JN (2016) Study on the relationship between Hepatitis C Virus Infection by different transmission route and genotypes in Guangxi. *Chinese General Practice* 19: 3534–3538.
25. Tang W, Su MH (2014) HCV genotyping distribution and epidemiological characteristics in Guangxi. *WCJD* 22: 1300–1306.
26. LV MY (2015) Establishment and preliminary application of methodology for HCV genotyping. [Master]: Chinese Center for Disease Control and Prevention.
27. Wang XM, Zhou J, Ye YX (2017) HCV genotyping and its clinical characteristics in HCV patients in west china hospital. *Chinese Journal of Evidence-Based Medicine* 17: 511–516.
28. Lin J, Wen XY, Tang M (2016) Characteristics and clinical Value of Hepatitis C Virus genotypes in South Sichuan. *Sichuan Medical Journal* 37: 1348–1351.
29. Xiao GR, Xu T (2016) Genotyping for 1998 HCV strains in west china hospital, 2009–2013. *Journal of Sichuan University(Medical Science Edition)* 47: 755–758.
30. Nie B, Zhang G, Guo Y, Li Q, Liu J, et al. (2015) Phylogenetic analysis of HCV subgenotypes in patients from Sichuan province in China based on the NS5B region. *Int J Mol Med* 36: 1028–1034. <https://doi.org/10.3892/ijmm.2015.2321> PMID: 26311523
31. Chen ZY, Xie N, Zhang ZX, Meng CR, Gu T, et al. (2016) HCV genotyping study among han, uygur in Xinjiang. *Chongqing Medical Journal* 45: 14–16, 18.
32. Wu YS, Bai JY, Yang CM (1997) Study on HCV genotyping in Urumqi. *Journal of Clinical Internal Medicine*: 265–266.
33. Li ZB, Jiang Y, Li H (2010) Hepatitis c genotype distribution research on 82 HCV infectors in Urumchi. *J Xinjiang Med Univ* 33: 542–544.
34. Chen SL, Wang YM, Li CM (2001) Molecular epidemiology of hepatitis C virus infection in intravenous drug addicts. *Hepatology*: 154.
35. Liu B, Wang YM (2002) HCV genotyping research on 69 cases in Chongqing. *Chongqing Medical Journal* 31: 174–175.
36. Fang YF, Chen S, Wang YM, Li CM (2002) Investigation on infection status of IDUs in Chongqing. *Chin J Drug Depend* 10: 220–222.
37. Fan K, Wang YM, Chen W. Investigation on genotype distribution, transmission pathway and discovery time of Chronic Hepatitis C Patients.; 2010; Guilin, Guangxi, China. pp. 1. PMID: 2010
38. Zhou ZG, Xu H (2015) HCV genotype distribution and its clinical significance in Chongqing. *Journal of Chongqing Medical University* 40: 383–387.
39. Zhao SM, Li M, Hu J (2006) HCV epidemiology investigation among unpaid blood donors in Chongqing. *Chongqing Medical Journal* 35: 964–966, 968.
40. Song SL, Gui WJ, Li XQ (2015) HCV genotype analysis among 169 HCV ingectors in Wuhan. *Practical journal of clinical medicine* 19: 36–38.
41. Li W, Zeng YL, Liu JP (2015) Distribution of HCV genotypes in patients with Chronic Hepatitis C in Henan. *Journal of Chinese Practical Diagnosis and Therapy*: 538–539.
42. Tang W (2014) Hepatitis c virus genotyping and its clinical significance in Guangxi [Master]. Guangxi: Guangxi Medical University.
43. Lang JY, Sun P, Lu S, Zhao PW, Wang CM (2017) HCV genotype distribution in Hohhot, Inner Mongolia. *Chinese Journal of Virology*: 61–66. PMID: 30702823
44. Tang ZY, Yang DL, Yu ZQ, Hao LJ (1995) The clinical manifestation of HCV infectors and the follow-up observation. *Shanghai Journal of Preventive Medicine*: 474–476.
45. Chen ZY (2015) HCV genotype investigation of 652 Han, Uygur and Kazak in Xinjiang [Master]: Xinjiang Medical University.

46. Song HB (2008) Anti -HCV reactivity of HCV infectors in different regions of China. *Chinese Journal of Public Health* 10.
47. Huang L, Ye DH, Ou XD (2018) HCV genotype distribution of the voluntary blood donors in Heyuan. *China Tropical Medicine*: 708–711.
48. Kaufman J, Jing J (2002) China and AIDS—the time to act is now. *Science* 296: 2339–2340. <https://doi.org/10.1126/science.1074479> PMID: 12089428
49. Shan H, Wang JX, Ren FR, Zhang YZ, Zhao HY, et al. (2002) Blood banking in China. *Lancet* 360: 1770–1775. [https://doi.org/10.1016/S0140-6736\(02\)11669-2](https://doi.org/10.1016/S0140-6736(02)11669-2) PMID: 12480443
50. Shi XL, Ren QH, Zhu ZY, Qu DM, Ji Y, et al. (1999) Hepatitis C virus infection in blood donors in the People's Republic of China. *Transfusion* 39: 913.
51. Ji YR, Q H. Zhu ZY. etal (1998) Investigation and analysis of HCV infection in Chinese blood donors. *Acta Acad Med Sinicae* 203: 240–241.
52. Lu L, An Y, Zou J, Gu L, Zhao Z, et al. (2015) The evolutionary patterns of hepatitis C virus subtype 2a and 6a isolates linked to an outbreak in China in 2012. *Virology* 485: 431–438. <https://doi.org/10.1016/j.virol.2015.08.012> PMID: 26343863
53. L L., W ZF., Q YM. (2014) Hepatitis C virus infection and its influencing factors in maintenance hemodialysis patients *Anhui Medical Journal* 35: 1045–1048.
54. Cui L (2015) Research on HCV genotype distribution among population being monitored with HIV/AIDS in Anhui [Master]. Anhui: Anhui Medical University.
55. Z L. (2012) Preliminary study on evolutionary origin of major HCV genotypes and construction of chimeric virus in Yunnan Province [Master]: Kunming University of Science and Technology.
56. Zhou DX, Tang JW, Chu IM, Cheung JL, Tang NL, et al. (2006) Hepatitis C virus genotype distribution among intravenous drug user and the general population in Hong Kong. *J Med Virol* 78: 574–581. <https://doi.org/10.1002/jmv.20578> PMID: 16555294
57. Aceijas C, Rhodes T (2007) Global estimates of prevalence of HCV infection among injecting drug users. *Int J Drug Policy* 18: 352–358. <https://doi.org/10.1016/j.drugpo.2007.04.004> PMID: 17854722
58. Garten RJ, Lai S, Zhang J, Liu W, Chen J, et al. (2004) Rapid transmission of hepatitis C virus among young injecting heroin users in Southern China. *Int J Epidemiol* 33: 182–188. <https://doi.org/10.1093/ije/dyh019> PMID: 15075167
59. Huang Y, Guo N, Yu Q, Lv Y, Ma H, et al. (2015) Risk factors for hepatitis B and C infection among blood donors in five Chinese blood centers. *Transfusion* 55: 388–394. <https://doi.org/10.1111/trf.12850> PMID: 25382751
60. Cruz-Rivera M, Carpio-Pedroza JC, Escobar-Gutierrez A, Lozano Dea (2013) Rapid hepatitis C virus divergence among chronically infected individuals. *J Clin Microbiol* 51: 629–632. <https://doi.org/10.1128/JCM.03042-12> PMID: 23224093
61. Fu Y, Qin W, Cao H, Xu R, Tan Y, et al. (2012) HCV 6a prevalence in Guangdong province had the origin from Vietnam and recent dissemination to other regions of China: phylogeographic analyses. *PLoS One* 7: e28006. <https://doi.org/10.1371/journal.pone.0028006> PMID: 22253686
62. (2012) The 14th International Symposium on Viral Hepatitis and Liver Disease (ISVHLD). *Chinese Journal of Gastroenterology and Hepatology* 21: 73.