Genotype and genetic variation of HCV infections with low-risk factors in Putian coastal regions, China

X. M. LI[1,2]†, R. X. QIU[3]†, C. H. SONG[4]†, Q. H. HUANG[3], X. D. WANG[5],
Z. T. HU[3], X. Z. HE[3], X. Y. YE[3], X. G. HUANG[3], F. F. ZHENG[3] AND G. X. LIN[3,6,7]*

[1] *Laboratory of Research Center, Affiliated Hospital (Group) of Putian University, Putian 351100, Fujian, China*
[2] *Medical College of Xiamen University, Xiamen 361102, Fujian, China*
[3] *Department of Infectious Diseases, Affiliated Hospital (Group) of Putian University, Putian 351100, Fujian, China*
[4] *Department of Gastroenterology, Affiliated Hospital (Group) of Putian University, Putian 351100, Fujian, China*
[5] *Department of Infectious Diseases, First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, Zhejiang, China*
[6] *Southern Medical University, Guangzhou 510515, Guangdong, China*
[7] *Fujian Medical University, Fuzhou 350117, Fujian, China*

SUMMARY

Hepatitis C virus (HCV) infection is one of the leading causes of death and morbidity associated with liver disease. Risk factors identified for the transmission of HCV include contaminated blood products, intravenous drug use, body piercing, an infected mother at birth, sexual activity, and dental therapy, among others. However, the exact diversity of the HCV genotype and genetic variation among patients with low-risk factors is still unknown. In this study, we briefly described and analysed the genotype distribution and genetic variation of HCV infections with low-risk factors using molecular biology techniques. The results suggested that genotype 1b was predominant, followed by genotypes 2a and 1a. Genetic variations in the 5′ UTR sequences of HCV were identified, including point mutations, deletions, and insertions. The frequency of genetic variations in 1b was higher than in 2a. This study provides considerable value for the prevention and treatment of liver disease caused by HCV among patients with low-risk factors and for the development of HCV diagnostic reagents and vaccines.

Key words: 5′ UTR, genetic variation, genotype, Hepatitis C virus, phylogenetic tree, risk factor.

INTRODUCTION

Hepatitis C virus (HCV) infection is one of the leading causes of death and morbidity associated with liver disease and is a major contributor to public health burdens worldwide [1, 2]. Recent estimates showed that approximately 3–4 million people around the world are newly infected every year, and over 350 000 patients die annually due to HCV-related disorders [3–5]. The prevalence rate of HCV infection in China is 0·43%, and approximately 13 million people are anti-HCV positive [6]. China is one of the most affected countries and currently has the largest number of HCV-infected people in the world [7].

Patients with chronic HCV infection are the most important source of transmission. Risk factors

* Author for correspondence: G. X. Lin, No. 999 East Dongzhen Road, Licheng District, Putian 351100, Fujian, China.
(Email: lingx609@163.com)
† These authors contributed equally to this work.

identified for the transmission of HCV include contaminated blood products, intravenous drug use, body piercing, an infected mother at birth, sexual behaviour, dental therapy, and other unclear factors [8–10]. A previous epidemiological investigation was conducted by the Department of Infectious Diseases at the Affiliated Hospital of Putian University from 2006 to 2013, revealing a high incidence (20·83%) of anti-HCV-positive patients among 3312 samples from villages along the coast of Putian. Surprisingly, the prevalence of HCV infection within these areas was extremely high compared with the average global level.

Research showed that the prevalence of HCV cases was contingent on country or region, with HCV cases representing <2% of the national population in developed countries and >15% in developing countries [11]. The prevalence of anti-HCV seroprevalence was lower than 1·5% in Asia Pacific, tropical Latin America, and North America [1]. The villages mentioned above are situated in the coastal areas of Putian in south-eastern China and are relatively impoverished and economically disadvantaged. Many patients seeking care in these villages suffer from liver problems.

In general, a high prevalence of HCV infection occurs among individuals or in areas where high-risk factors are common. However, many HCV-infected patients from the coastal rural areas of Putian were not exposed to known hazardous environments in this epidemiological investigation. In addition, we found a distinct phenomenon of familial aggregation in these areas. For example, two or even six members in some families (21%) were positive for HCV infection.

It is important to know the mechanism of transmission and pathogenesis to enable appropriate treatment and prevention of HCV infection. Indeed, the familial aggregation of HCV infection provided us with a miniature model of the pathogenesis of HCV infection. Characteristics and changes of biological genes contribute to the solution to this problem [12]. Additionally, HCV exhibits enormous genetic diversity and a high degree of genetic similarity, which can be used to determine the individual HCV isolates of various subtypes [13]. Genotyping is useful for investigating infection outbreaks and for improving the general understanding of the epidemiological and virological features of HCV infection [14, 15]. Petruzziello et al. presented an epidemiological update of HCV genotype distribution across Europe and found a prevalence of 1·7% from 2000 to 2015.

Among those cases, the frequency of genotype 1 was 55·1–70·0%, followed by genotype 2 (3·2–8·9%), genotype 3 (21·0–29·0%) and genotype 4 (4·9–5·8%) from 2000 to 2015 [3]. They also studied the global HCV epidemiology using available data published between 2000 and 2015 from 138 countries (approximately 90% of the global population), revealing that genotypes 1 (49·1%) and 3 (17·9%) were common worldwide, followed by genotypes 4 (16·8%) and 2 (11·1%), while genotypes 5 and 6 together represented <5% [16]. However, the precise diversity of HCV genotypes and genetic variation among patients with low-risk factors was still unknown.

The HCV genome contains a 5′ untranslatable region (5′ UTR), a highly conserved RNA structure with a single open reading frame encoding structural proteins (Core, E1, E2) and non-structural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B), and a 3′ UTR [17]. HCV genotyping assays are currently based on the sequence of regions (5′ UTR, Core, E1, and NS5B) [15, 18]. Sequencing of an appropriate region, such as the 5′ UTR, core, E1, or NS5B, is the gold standard for discriminating HCV genotype. Furthermore, the 5′ UTR has been widely utilized for genotype distribution due to its high degree of conservation [19].

In this study, we combined the molecular biology techniques with genotype distribution, genetic variability, and phylogenetic analysis to further identify the genotype and genetic variation of HCV-infected patients with low-risk factors in Putian coastal regions of China. A potential genetic mechanism of transmission and pathogenesis in these individuals with a low risk of HCV infection might provide important information for future treatment and prevention of HCV infections.

MATERIALS AND METHODS

Ethics statement

Ethical approval for the study was obtained from the Affiliated Hospital of Putian University, China. Informed consent was obtained from each patient before the specimens and survey information were collected.

Subjects and inclusion criteria

Samples were collected from the coastal rural areas of Putian City in south-eastern China, in which

participants had anti-HCV antibodies or tested positive for HCV RNA. The patients had never received treatment for chronic infection with HCV. The expression of HCV RNA in each patient was more than 10^3 copies/ml. Blood samples were gathered and serum was isolated according to conventional methods. All eligible serum samples had detectable HCV RNA obtained from patients who were anti-HCV positive according to their clinical diagnosis. Additionally, a complete and detailed epidemiological questionnaire was used to assess the risk factors for HCV infection, including age at diagnosis, gender, ALT (alanine aminotransferase) levels, educational level, history of liver disease, family medical history, blood donation/transfusion history, dental therapy and surgical intervention. Valid follow-up information for the participants was recorded between July 2014 and December 2016. Diagnostic criteria for HCV infection were based on the new prevention guide on chronic hepatitis C infection in China, 2015 [20].

Exclusion criteria

All patients were negative for HBV (hepatitis B virus), HDV (hepatitis D virus), and HIV (human immunodeficiency virus) antibodies. Patients with known alcohol abuse, drug problems, pregnancy, and Child-Pugh grade B/C were also not considered in this study. Patients who had high-risk factors, such as a medical history of selling blood, blood transfusion, receiving blood products, haemodialysis, intravenous drug use, body piercing, an infected mother at birth, high-risk sexual behaviour and dental therapy, were excluded from this study.

Verification method

Plasma was preserved at −70 °C until further use or at −20 °C for 4 h for immediate use. The samples were measured for anti-HCV antibodies using commercial third-generation enzyme immunoassays from two manufacturers (InTec Products Core Anti-HCV EIA, C 20 100 068; Kehua Core Anti-HCV EIA, C2101012061, Shanghai, China). The serum samples that reacted with both of the kits were determined to be positive for HCV infection, although autoimmunity disorders can sometimes generate false-positive results. The method used to verify the detection of HCV RNA can increase the reliability of the results. The participants were labelled according to a specific numbering system to identify their individual information.

RNA extraction and amplification

HCV RNA was extracted from the serum of subjects using the QIAmp Viral RNA Kit according to the manufacturer's instructions (Qiagen, Chadstone Centre, Vic, Australia). The extracted RNA was eluted in 50–100 µl RNase-free buffers or distilled water. Complementary DNA (cDNA) was synthesized from 5 µl of eluted RNA using the SuperScript VILO cDNA Synthesis Kit (Life Technologies, Carlsbad, CA, USA) with random hexamer primers. The reactions were heated on a thermocycler (Roche Applied Sciences, Mannheim, Germany) at 37 °C for 15 min, 50 °C for 5 min, and 95 °C for 5 min and maintained at 4 °C [21].

The HCV genotype was determined via nested-polymerase chain reaction (PCR) and sequencing based on the 5′ UTR region of the genome. The primers used for the nested-PCR amplification of the fragment of the 5′ UTR region are listed in Table 1. A product of approximately 230 bp from the 5′ UTR region was amplified with a modified version of a previously described protocol [13]. Then, the 20 µl reaction mixture (containing 5 µl template cDNA) was reacted in the thermocycler. The first round to amplify a 340 bp fragment was performed for 40 cycles using the outer primers Out-F and Out-R with initial denaturation at 94 °C for 5 min, denaturation at 94 °C for 10 s, annealing at 56 °C for 20 s, and elongation at 72 °C for 1 min, followed by a final extension at 72 °C for 5 min [22]. Then, 2 µl of PCR product was used for the second-round amplification of a 230 bp fragment in a 20 µl reaction mixture containing the inner primers In-F and In-R, and the thermal cycling parameters described above were applied. Subsequently, 3 µl of PCR product was electrophoresed on a 1% agarose gel with the DL2000 marker. The amplified products were quantified using a NanoDrop spectrophotometer.

Sequencing and phylogenetic analysis

The amplicons were purified using the PCR Product Purification Kit (Roche Diagnostics GmbH, Mannheim, Germany) for sequencing to identify the HCV genotype. DNA sequencing was performed by two different companies: Adicon Clinical Laboratories, Inc., and KingMed Clinical Laboratory, China. The

Table 1. *Outer and inner primers for nested-PCR amplification of the 5′ UTR region from HCV genomic RNA*

Genomic region	Name	Sequence of primers[a]	Position[b]
5′ UTR	Out-F	5′-GGGCSACACTCSACCATR-3′	2–28
	Out-R	5′-GTGNGGTTTAGGATTCRTGCT-3′	304–330
	In-F	5′-HATCACTCCCCTGTGAGGAACTW-3′	35–68
	In-R	5′-GTCTAGCCATGGCGTTAGTA-3′	227–289

[a] For degenerate primers, R = A or T, W = A or T, S = G or C, K = G or T, Y = C or T, B = C or G or T, D = A or G or T, H = A or C or T, V = A or C or G, N = A or C or G or T.
[b] Position with reference to HCV isolate sP256432, GenBank accession KU180718.1.

sequencing chromatogram was visualized using Chromas v1·45chs. The assembled sequences were analysed using Lasergene v8 (DNASTAR, Madison, WI, USA). The HCV genotype was assigned using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) or the HCV Sequence database (http://hcv.lanl.gov/content/sequence/HCV/ToolsOutline.html) [23]. The sequences of the 5′ UTR fragment were aligned with reference sequences from the GenBank database (https://www.ncbi.nlm.nih.gov/nucleotide) using ClustalX v1·8·31. A phylogenetic tree was constructed according to the neighbour-joining method using DIVEIN (PhyML 3·0) and MEGA 4 [15].

Data and statistical analysis

The experiments were carried out with a minimum of two replicates. Means and standard deviations (s.D.s) were calculated and analysed for continuous variables. One-way ANOVA was used to determine the differences among the variables. The Chi-square test or Fisher's exact test was used to identify differences in the frequencies of genotype distribution and genetic variation. All statistical analyses were performed using statistics software (Stata version 12·0, StataCorp, College Station, TX). A probability (P) value <0·05 was considered to be statistically significant.

RESULTS

Epidemiological characteristics of HCV patients

According to the criteria listed above, a total of 254 full and detailed cases that included blood serum and questionnaires were eventually included in this study. The clinicopathological and demographic characteristics of 254 HCV-infected patients from coastal rural areas of Putian City in south-eastern China are shown in Table 3. Among them, 101 (39·8%) were males and 153 (60·2%) were females. Overall, the HCV-positive patients selected for this study were 21–82 years old (mean age of 55 ± 34 years). The prevalence of HCV infection was approximately 2% in the <30 years age group, 35% in the 30–50 years age group, and 63% in the >50 years age group. People >50 years of age had a significantly greater frequency of HCV infection. Furthermore, we found that several HCV patients had infected parents, brothers or sisters, such as isolate No. 138-31469 (1b), who had a daughter with an HCV infection; isolate 180-31589 (2a), whose elder brother and sister were infected with HCV; and isolate 185-31602 (1b), who had an older sister with an HCV infection. All of the patients were of Han ethnicity.

Amplicon analysis

HCV-infected samples were further assessed for genetic variability by sequencing the 5′ UTR region, conducting a BLAST analysis, and determining the genotype distribution. The 5′ UTR region of the HCV genome (from nucleotide (nt) 2 to 340) was amplified from the total RNA of 254 HCV-positive samples collected in the villages of Putian coastal areas. The amplification of the 5′ UTR fragment was performed using nested-PCR, and the primers are shown in Table 1. A band of approximately 230 bp was detected using gel electrophoresis to confirm the size and quality of the band. Then, the fragments were independently sequenced by two companies. The electrophoresis results and chromatograms of the obtained sequences in the 5′ UTR region of HCV isolates are shown in Figures 1 and 2, respectively.

BLAST and sequence analysis

The 5′ UTR sequencing results from two independent companies had a high similarity of 99·8%. One of the HCV samples that were analysed in a BLAST
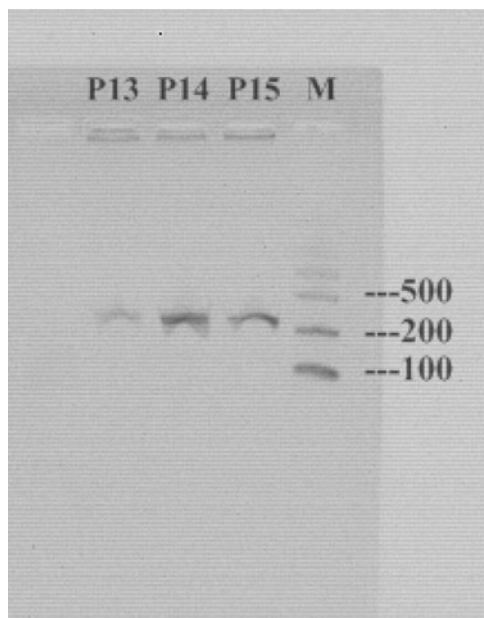
Fig. 1. Gel electrophoresis results for 5′ UTR fragments amplified by nested-PCR. M: DL2000 marker; P13: HCV isolate No. 013-30635.

homology search was listed after a 100% analogous identity was found. The nucleotide sequence of the 5′ UTR fragment from HCV-positive isolate No. 013-30635 was determined: TCTTGCGGGGG CACGCCCAAATGGCCGGGCATAGAGTGGGGT TTATCCAAGAAAGGACCCAGTCTTCCCGGCA ATTCCGGTGTACTCACCGGTTCCGCAGACCA CTATGGCTCTCCCGGGAGGGGGGGGGCCTGG AGGCTGTACGACACTCATACTAACGCCATGG CTAGACGCTTTCTGCGTGAAGACAGTAGTTC CTCACAGGGGAGTGA.

Nucleotide sequences of the 5′ UTR regions from 254 HCV RNA-positive samples were used as the input query for NCBI BLAST to find the most likely similar sequences. BLAST results revealed that highly similar sequences included identification of the 5′ UTR regions of the HCV genome. Details of the similarity of partial BLAST analysis are summarized in Table 2. The overall per cent nucleotide identity (PNI) in the 5′ UTR regions of the HCV genome was 97–100%. A partial BLAST result of the No. 002-18325 isolate with 100% of PNI and 414 of the maximum score is shown in Table 2. BLAST analysis revealed that the amplification of the 230 bp fragment was 100% homologous with the hepatitis C virus isolate sP256432 polyprotein gene (accession number KU180718·1) from nt 37 to nt 260.

The multiple sequence alignment using ClustalX software v1·8·31 for the 254 HCV RNA-positive samples compared with the GenBank database is presented in Figure 3. The nucleotide sequences of the 5′ UTR region were considered to be highly conserved. However, some mutations and variations in the 5′ UTR region of HCV gene are shown in Figure 3.

Genotype distribution of the patients

Four HCV genotypes (1, 2, 3, and 6) and subtypes were detected among the 254 HCV-infected patients via sequencing analysis. The majority of the HCV-positive samples belonged to genotype 1b (59·8% 152/254), followed by genotypes 2a (30·7% 78/254), 1a (5·1% 13/254), 1c (2·0% 6/254), 6b (0·8% 3/254), and 2d (0·4% 2/254), as shown in Table 3. However, no other genotypes were detected.

The genotype distribution of HCV-positive patients is presented in Figure 4. As shown in Table 3 and Figure 4, the overall genotype ratios of HCV-positive females were higher than those of males. Samples with genotype 1a (average age of 56 ± 12 years) included six (46·2%) males and seven (53·8%) females. Of the genotype 1b patients with a mean age of 55 ± 34 years, 66 (43·4%) were males and 86 (56·6%) were females. Of the 78 patients with genotype 2a (mean age of 54 ± 26 years), 23 (29·5%) and 55 (70·5%) were males and females, respectively. There was no statistically significant difference ($P > 0·05$) in the average age among HCV patients infected with genotypes 1a, 1b, and 2a.

The youngest HCV-positive patient was a 21-year-old male with genotype 1b. The oldest patient was an 82-year-old female with genotype 1b. The correlation between age and genotype (including 1a, 1b, and 2a) distribution is illustrated in Figure 5. In this study, no significant differences were observed in the age-specific prevalence of different HCV genotypes (1a, 1b, and 2a). The majority of HCV-infected patients were >50-years old (61·5% of genotype 1a, 65·8% of genotype 1b, and 62·8% of genotype 2a). Low prevalence rates for patients in the age range of 30–50 years were found, with 38·5% presenting genotype 1a, 32·2% presenting genotype 1b, and 34·6% presenting genotype 2a.

Genetic mutations and variations

Sequences in the 5′ UTR region of 254 HCV infections were analysed for genetic mutations, including insertions, deletions, transitions, and transversions. High mutation rates (18·1%, 46/254) were identified
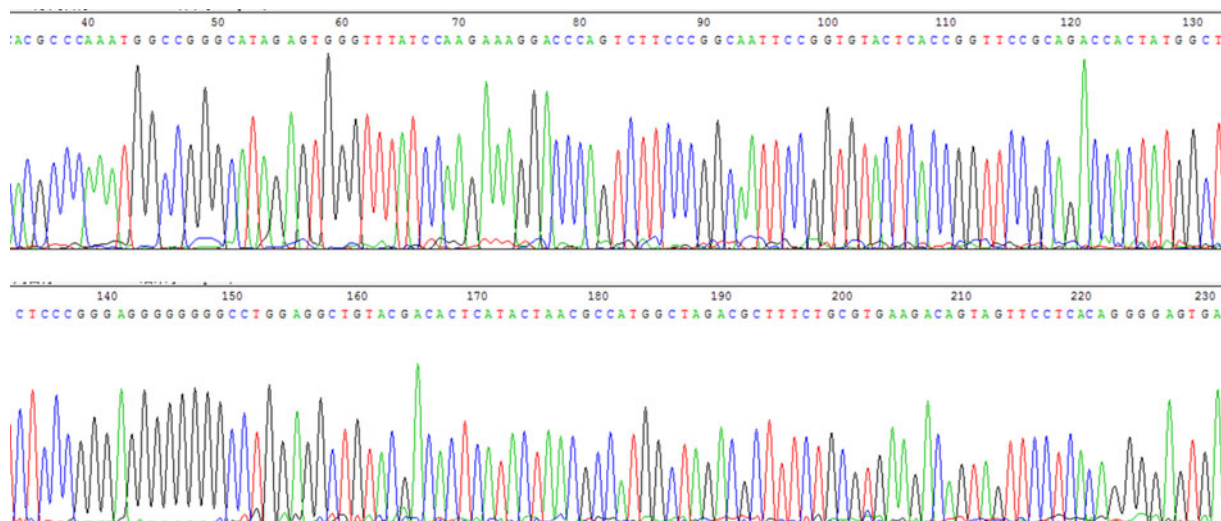
**Fig. 2.** Chromatogram of the sequence obtained from the 5′ UTR region of HCV isolate No. 013-30635.

Table 2. *NCBI nucleotide BLAST results for the sequence in the 5′ UTR region of HCV isolate No. 002-18325 (partial alignment of the results)*

Description	Max score	E-value	Identity, 100%	Accession
Hepatitis C virus isolate sP256432 polyprotein gene	414	3e-112	100	KU180718·1
Hepatitis C virus strain TSBD7775 5′ UTR	414	3e-112	100	KX084739·1
Hepatitis C virus subtype 1b genomic RNA	414	3e-112	100	LC011931·1
Hepatitis C virus isolate TV503 polyprotein gene	414	3e-112	100	KM252781·1
Hepatitis C virus subtype 1b genomic RNA	414	3e-112	100	AB828699·1
Hepatitis C virus subtype 1b	414	3e-112	100	GU133617·1
Hepatitis C virus subtype 1b isolate LB	414	3e-112	100	HQ110091·1
Hepatitis C virus isolate NBITHC04 5′ UTR	414	3e-112	100	JN387057·1

Table 3. *Correlation between clinicopathological features and HCV genotype frequencies among 254 HCV-infected patients from the Affiliated Hospital of Putian University, China*

| Clinicopathological feature	HCV genotype						
	Genotyoe 1 (%)			Genotype 2 (%)		Genotype 3 (%)	Genotype 6 (%)
	1a	1b	1c	2a	2d	3b	6b
Sex							
M	6 (46·2%)	66 (43·4%)	2 (40·0%)	23 (29·5%)	1	1 (33·3%)	2
F	7 (53·8%)	86 (56·6%)	3 (60·0%)	55 (70·5%)	0	2 (66·7%)	0
Age (year)							
<30	0	3 (2·0%)	0	2 (2·6%)	0	0	0
30–50	5 (38·5%)	49 (32·2%)	4 (80·0%)	27 (34·6%)	0	2 (66·7%)	2
>50	8 (61·5%)	100 (65·8%)	1 (20·0%)	49 (62·8%)	1	1 (33·3%)	0
Age Mean ± s.d.	56 ± 12	55 ± 34	48 ± 3	54 ± 26	64	48 ± 12	38 ± 7
Total	13 (5·1%)	152 (59·8%)	5 (2·0%)	78 (30·7%)	1 (0·4%)	3 (1·2%)	2 (0·8%)
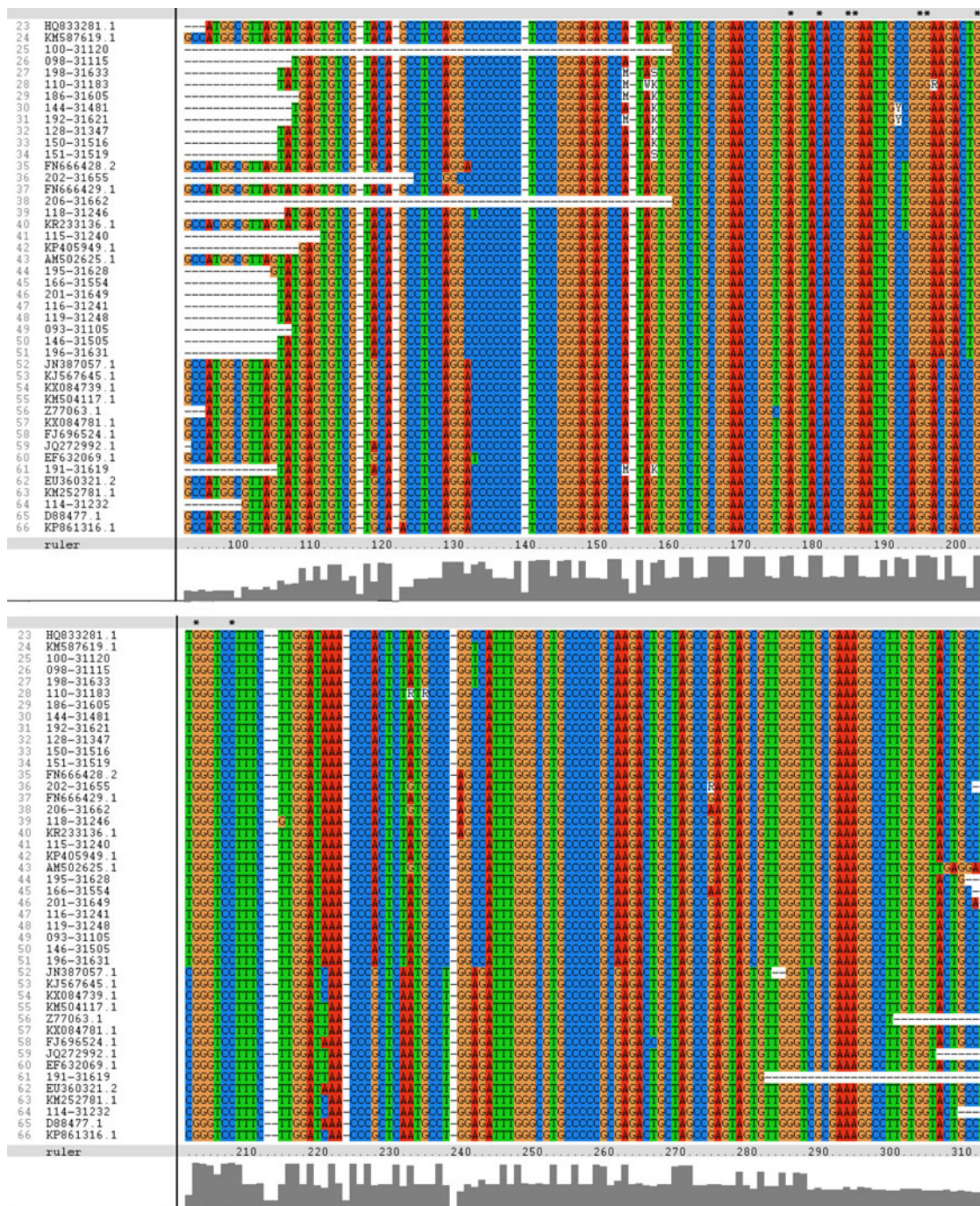
Fig. 3. Comparison of nucleotide sequences of the partial 5′ UTR region from 254 HCV RNA-positive patients and reference sequences obtained from the GenBank database. The partial multiple sequence alignment was produced by ClustalX v1·8·31 software.

in the 5′ UTR sequences, with 58·7% (27/46) of mutations found in genotype 1b and 32·6% (15/46) found in genotype 2a. Point mutations (i.e., base substitutions) were identified in 14·2% (36/254) of the sequences, and 5·1% (13/254) of specimens contained base deletions and insertions. Among these point mutations, all

samples had a single point mutation in the 5′ UTR region, dual point mutations in four specimens, and three point mutations in two specimens. The majority of point mutations in the 5′ UTR region appeared ranging from nt −122 to nt −264. Single point mutations of the 5′ UTR region occurred with high frequency at
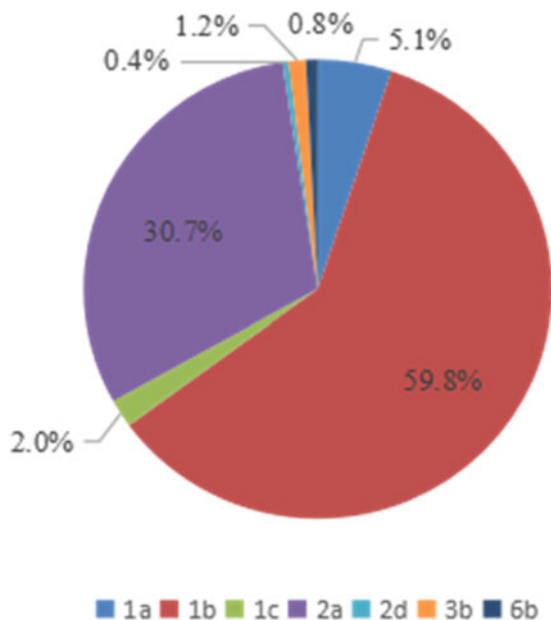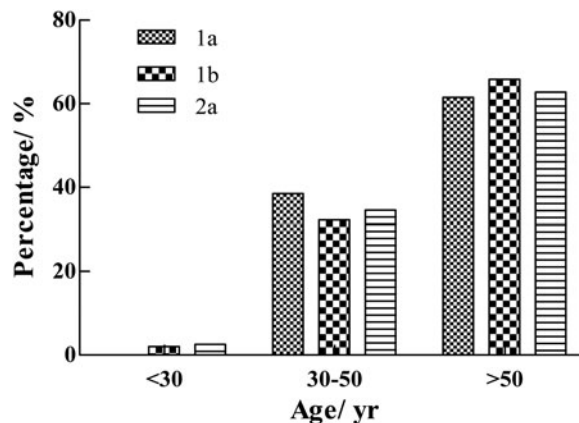
Fig. 4. Genotype distribution of 254 HCV-positive patients from the Affiliated Hospital of Putian University, China.

nt −86 ($n = 2$), nt −94 ($n = 2$), nt −127 ($n = 2$), nt −138 ($n = 2$), nt −139 ($n = 2$), nt −200 ($n = 2$), nt −221 ($n = 2$), and nt −264 ($n = 2$). A sample showed three point mutations in the 5′ UTR region at nt −127, nt −224, and nt −226. The base insertion of HCV samples ($n = 6$) was found in the 5′ UTR region within a range of nt −202 to nt −237, and two of these insertions presented at the same site (nt −221). The deletions in the 5′ UTR region of HCV samples ($n = 7$) were discovered within a range of −83 to −294 nucleotides in length, and three of those deletions presented at the same site (nt −294). Transitions and transversions failed to be detected. Interestingly, there was a particular sample that exhibited two deletions, at nt −91 and nt −114, and one point mutation at nt −182.

Phylogenetic analysis

The high-scoring reference sequences from the GenBank database were homologous to the query sequence of the 5′ UTR region from 254 HCV-infected isolates. The phylogenetic tree was constructed with the neighbour-joining method (PhyML 3·0 and MEGA 4) using 500 bootstrap replicates for sequence alignments, as shown in Figure 6. According to the phylogenetic tree, the amplified fragments of 254 HCV RNA-positive patients were analysed along with 111 reference sequences of distinct


Fig. 5. Age-specific prevalence of genotypes (1a, 1b, and 2a) of HCV-positive samples from the Affiliated Hospital of Putian University, China.

genotypes derived from NCBI BLAST results. Phylogenetic analysis revealed that the main genotypes in 254 HCV-positive patients from the coast of Putian were genotypes 1b and 2a. Four distinct genotypes (1, 2, 3, and 6) were identified in this study. For example, genotype 1b isolate No. 009-30622 was highly homologous to the reference strain sP256432, which was reported by Simmonds and Barnes from University of Edinburgh, UK, in 2015. Comparatively, the genotype 2a isolate No. 017-30643 was strongly consistent with MRS40 at 100% homology, which was submitted by Jordier in 2012 in France. Therefore, the phylogenetic tree of HCV 5′ UTR sequences demonstrated that 152 (59·8%) strains were closely related to subtype 1b and 78 (30·7%) strains were of subtype 2a. The prevalences of other subtypes are illustrated in Figure 6 and Table 3. Furthermore, the sequences in the 5′ UTR region of HCV isolates discovered in this study will be uploaded onto the NCBI platform.

DISCUSSION

The genotypes and subtypes of HCV vary greatly in different areas, and the distributions of these genotypes have obvious geographical and population differences [24–27]. Genotyping is beneficial for analysing the route of transmission of hepatitis C and is helpful for the early detection of new subtypes of epidemic outbreaks that may have serious consequences and for providing a scientific basis for the prevention and control of HCV [28, 29]. With the rapid development of globalization and increasing linkage among the international community, the international
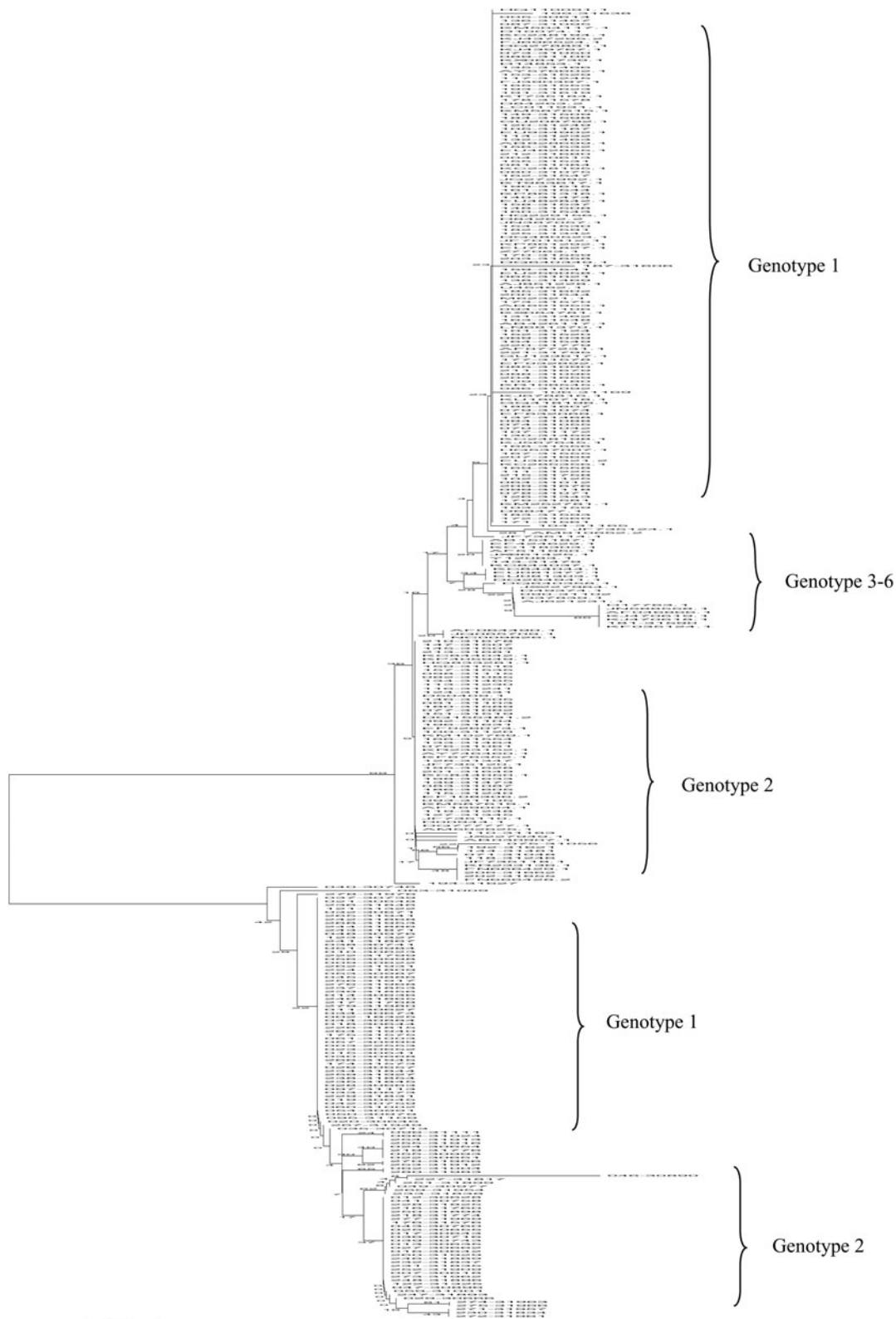
Fig. 6. Phylogenetic analysis of the 5′ UTR region from 254 HCV-infected isolates and reference sequences were generated via the neighbour-joining method using DIVEIN (PhyML 3·0) and MEGA 4 with 500 bootstrap replicates. Sequences of HCV-positive samples from the Affiliated Hospital of Putian University, China, were labelled as No. 163-31548. The reference sequences are listed by their GenBank accession number. The branch lengths were proportional to the evolutionary distance between the sequences and distance scale.

population flow has become more frequent [30]. This phenomenon has led to the loss of geographical barriers, resulting in the rapid transmission of the HCV genotype across countries. Therefore, the regional distribution of the HCV genotype provides support for epidemiological studies.

HCV genotype 1 is found worldwide, and it is the major genotype found in the USA, Europe, Australia and Japan [19, 31]. In most parts of China, the predominant HCV genotype is genotype 1b, followed by genotype 2a [32, 33]. However, there are differences in genotypes among different regions. The most prominent feature of HCV in China is that there is a low prevalence rate in most areas but a high prevalence rate in some regions along with a scattered and uneven distribution of cases. Our previous epidemiological investigation found that some villages have extremely high HCV infection rates compared with the average global level. These villages are found in the coastal areas of Putian in southeastern China, and many of these patients have low-risk factors for HCV infection. However, the exact diversity of HCV genotype and genetic variation in these patients with low-risk factors had remained unknown. In the present study, we investigated the HCV genotype by sequencing the 5′ UTR region of 254 HCV-infected positive samples with low-risk factors from the Affiliated Hospital of Putian University (China) to analyse the genotype distribution and genetic variation. According to Table 3, genotypes 1b (59·8%) and 2a (30·7%) were the most prevalent genotypes, followed by genotype 1a (5·1%). Additionally, genotypes 3b and 6b were detected in only 1·2% and 0·8% of cases, respectively. Genotypes 4, 5, and 7 were not detected in our study, which was an outcome similar to the results of Dunford et al. [12]. The distribution of genotypes and subtypes varies according to geography. Our results were also consistent with the results of Yin et al., who found 36 samples (53·73%) that were subtype 1b and 31 samples (46·27%) that were subtype 2a in rural Hebei Province [15]. However, some differences between the present study and that of Dunford et al., who reported subtypes 1a (33%), 1b (27%), 2a (0·4%), 3a (0·7%), 3b (1·1%), 6a (18·8%), 6e (6·0%), 6 h (4·6%), and 6 l (6·4%) of HCV infections in Vietnam [12]. In addition, gender distribution showed that the prevalence rates of major subtypes (1a, 1b, 1c, 2a, and 3b) of anti-HCV-positive females were higher than in males (Table 3). However, Yin et al. reported that HCV subtype 2a was predominant in females and

subtype 1b was predominant in males [15]. This feature indicated a discrepancy in gender distribution based on country and region.

HCV infections are a major source of transmission. Risk factors identified for the transmission of HCV include transfusion of blood products, intravenous drug use (IDU), body piercing, an infected mother at birth, high-risk sexual behaviour, and other unconfirmed methods [10, 34, 35]. Astoundingly, the HCV-infected patients from the coastal rural areas of Putian included in this survey do not have exposure to known high-risk factors. In addition, with the implementation of medical hygiene procedures and adequate screening of blood donors over the past 10 years, the blood source and iatrogenic routes of transmission have been restricted. Nevertheless, transmission routes in areas with a high prevalence (20·83%) of anti-HCV positivity and the significant familial aggregation phenomena in the rural coastal regions of Putian remain an enigma. Contemporary studies have shown that high incidences of HCV infection have primarily occurred in undeveloped countries and regions, such as Africa and Southeast Asia. For example, the high prevalence of HCV positivity in Egypt was >10% of the general population, and this prevalence rose up to 16·8% among the blood donor population [11, 36]. Hanafiah et al. also reported in 2005 that the prevalence of HCV infection varied depending on country and region, such as Central and East Asia (>3·5%), Australasia and Western Europe (1·5–3·5%), and North America (<1·5%) [1]. Huang et al. reported that the high prevalence (28·5%) of anti-HCV positivity was due to blood selling in the late 1980s in rural Hebei Province, North China [37]. The prevalence of HCV-infected individuals with low-risk factors (including voluntary blood donors and pregnant women) was only 0·5%, and the prevalence rates were significantly higher among intravenous drug users (55·6%) and dialysis patients (26·6%) from Vietnam [12]. Previously, a high prevalence rate (28·86%, 303/1050) of the anti-HCV-positive cases in a village of Putian was reported by Lin et al. [38]. Furthermore, we speculated that the high prevalence (20·83%) of anti-HCV positivity and the significant familial aggregation phenomena might be caused by the HCV lineages inherited from ancestors. However, we did not consider vertical transmission to be an important factor because the incidence rate of HCV infection in individuals <30 years (20–30 years) was approximately only 2%, with a rate of 35% in individuals in the

30–50-year-old age group and 63% in the >50 year group, which was the main infection group. Although the patients did not present any history of blood selling, blood transfusion, drug abuse, hemodialysis, body piercing, vertical transmission, sexual behaviour or dental therapy in our study, the ancestors of these patients may have been be infected with HCV through some unknown route. In addition, we predicted that individuals might have been infected through vaccine inoculation for immunization in the 1980s when the safety and preventive treatments for iatrogenic transmission were not effective. In addition, the high prevalence of HCV infection in the coastal rural areas of Putian City may be due to the absence of effective prophylactic treatments among family members to prevent chronic HCV infection and patient-to-patient transfusion.

With respect to genetic evolution, we also analysed the genetic variation for these patients using molecular biology techniques. The results demonstrated that genetic variation was present in 18·1% (46/254) of the 5′ UTR sequences. Meanwhile, the proportion of mutation and variation in genotype 1b was 58·7% (27/46), which was more than in genotype 2a (32·6%, 15/46). Genetic variations in the 5′ UTR sequences of HCV were identified, including point mutations, deletions, and insertions. Basically, the results were consistent with most studies, and no transitions or transversions were detected. The genetic variability of the virus is an important factor for the limitation of clinical treatment. Kanwal and Mahmood identified the genetic variability and modification in the genomic regions of 5′ UTR, the core and NS5B as the underlying factors enabling the virus to escape the immune response [39]. González-Candelas et al. also reported that viral genetic variability might play an important role in the HCV transmission pathway, host susceptibility factors for infection or immune system response and serve as a powerful tool for tracing transmission [18]. Therefore, we can speculate that the genetic diversity and variability in the 5′ UTR fragment might be the reason HCV is able to evade host defences and develop drug resistance.

There are a few limitations in this study. Relevant research methods were mainly used to analyse the gene sequence of the 5′ UTR fragment but did not include the core and NS5B gene sequences. In addition, some patients with HCV may experience repeated infections. Therefore, the ultimate reason for high prevalence of patients with low-risk factors

is still unknown. We suggest that the route of HCV infection in these areas may be transmitted by lineages inherited from the ancestors. Additionally, HCV infection may have occurred via iatrogenic factors of vaccine inoculation before the 1980s, when China was relatively impoverished and economically undeveloped. Thus, a lack of adequate prophylactic treatments may be a general cause for patient-to-patient transfusion, which should be further investigated in subsequent studies.

CONCLUSION

In this study, we analysed the genotype distribution and genetic variation of HCV-infected patients with low-risk factors using molecular biology techniques. The results showed that genotype 1b was the most common, followed by genotype 2a and genotype 1a. In addition, genetic variations in the 5′ UTR sequence of HCV were identified, including point mutations, deletions, and insertions. The frequency of genetic variation in genotype 1b was higher than in genotype 2a. We suggest that the genetic variability of HCV may play an important role in immunosurveillance and clinical treatment for abnormal binding sites. Therefore, genetic variation may have an important effect on antiviral therapy for HCV. These variations may also provide an important experimental basis for the diagnosis and the treatment of liver disease caused by HCV in these coastal regions. Meanwhile, it has practical value for the development of diagnostic reagents and HCV vaccines.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

G. X. LIN conceived and designed the experiments; X. M. LI, R. X. QIU, and C. H. SONG performed

the surveys; X. M. LI and C. H. SONG analysed the data; Q. H. HUANG, X. D. WANG, Z. T. HU, X. Z. HE, X. Y. YE, X. G. HUANG, and F. F. ZHENG contributed materials; X. M. LI, R. X. QIU, and C. H. SONG wrote the paper.

DECLARATION OF INTEREST

The authors declare no conflicts of interest.

INFORMED CONSENT

Written informed consent was obtained from the patients prior to study participation.

REFERENCES

1. **Hanafiah KM, Groeger J, Flaxman AD.** Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* 2013; **57**(4): 1333–1342.
2. **Iqbal J, Mcrae S, Mai T.** Role of hepatitis C virus induced osteopontin in epithelial to mesenchymal transition, migration and invasion of hepatocytes. *PLoS ONE* 2014; **9**(1): e87464.
3. **Petruzziello A, et al.** Hepatitis C virus (HCV) genotypes distribution: an epidemiological up-date in Europe. *Infectious Agents and Cancer* 2016; **11**: 53–62.
4. **Perz JF, et al.** The contributions of hepatitis bvirus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *Journal of Hepatology* 2006; **45**: 529–538.
5. **Hanafiah KM, et al.** Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* 2013; **57**: 1333–1342.
6. **Chen YS, Li L, Cui FQ.** A sero-epidemiological study on hepatitis C in China. *Chinese Journal of Epidemiology* 2011; **32**: 888–891.
7. **Wu HB, Zhou ZX, Huang YX.** Analysis of epidemiological characteristics of viral hepatitis C in China, 2004–2011. *Modern Preventive Medicine* 2015; **7**(42): 1173–1175.
8. **Araujo AC.** Antibody-and genome-based identification of recent HCV infection. *Antiviral Therapy* 2012; **17**(7): 1459–1464.
9. **Rossi LMG, Rahal P.** Challenges in molecular epidemiology of hepatitis C virus. *Journal of Clinical Virology* 2014; **60**(2): 174–176.
10. **Simmonds P.** The origin of hepatitis C virus. *Current Topics in Microbiology and Immunology* 2013; **369**: 1–15.
11. **Hajarizadeh B, Grebely J, Dore GJ.** Epidemiology and natural history of HCV infection. *Nature Reviews Gastroenterology and Hepatology* 2013; **10**: 553–562.
12. **Dunford L, et al.** Hepatitis C virus in Vietnam: high prevalence of infection in dialysis and multi-transfused patients involving diverse and novel virus variants. *PLoS ONE* 2012; **7**(8): e41266.
13. **Chueca N, Rivadulla I, Lovatti R.** Using NS5B sequencing for hepatitis C virus genotyping reveals discordances with commercial platforms. *Plos ONE* 2016; **11**(4): e0153754.
14. **Li D, et al.** Epidemiology of hepatitis C virus infection in highly endemic HBV areas in China. *PLoS ONE* 2013; **8**(1): e54815.
15. **Yin WJ, et al.** Risk factors of hepatitis C virus transmission and genotype distribution in former blood donors from Chinese rural area. *BMC Public Health* 2015; **15**: 184–192.
16. **Petruzziello A, Marigliano S, Loquercio G.** Global epidemiology of hepatitis C virus infection: an up-date of the distribution and circulation of hepatitis C virus genotypes. *World Journal of Gastroenterology* 2016; **22**(34): 7824–7840.
17. **Moradpour D, Penin F.** Replication of hepatitis C virus. *Nature Reviews Microbiology* 2007; **5**: 453–463.
18. **González-Candelas F, Guiral S, Carbó R.** Patient-to-patient transmission of hepatitis C virus (HCV) during colonoscopy diagnosis. *Virology Journal* 2010; **7**: 217–226.
19. **Marco LD, et al.** Prevalence and follow-up of occult HCV infection in an Italian population free of clinically detectable infectious liver disease. *PLoS ONE* 2012; **7**(8): e43541.
20. **Chinese Society of Hepatology (CMA), Society of Infectious Diseases (CMA).** Guidelines for the prevention and treatment of hepatitis C. *Chinese Journal of Hepatology* 2015; **7**(3): 19–31.
21. **Lamoury FMJ, Jacka B, Bartlett S.** The influence of hepatitis C virus genetic region on phylogenetic clustering analysis. *PLoS ONE* 2015; **10**(7): e0131437.
22. **Makvandi K, Ranjbari N, Makvandi M.** Study of the association of mutant HBsAg gene and Hodgkin and NonHodgkin lymphoma. *Jundishapur Journal of Microbiology* 2015; **8**(11): e25726.
23. **Wasitthankasem R, et al.** Genotypic distribution of hepatitis C virus in Thailand and Southeast Asia. *Plos ONE* 2015; **10**(5): e0126764.
24. **Nguyen MH, Keeffe EB.** Chronic hepatitis C: genotypes 4 to 9. *Clinics in Liver Disease* 2005; **9**(3): 411.
25. **Jiao Y, et al.** Hepatitis C virus subtype and evolution characteristic among drug users, men who have sex with men, and the general population in Beijing, China. *Medicine (Baltimore)* 2016; **95**(6): e2688.
26. **Daw MA, El-Bouzedi A, Dau AA.** Geographic distribution of HCV genotypes in Libya and analysis of risk factors involved in their transmission. *BMC Research Notes* 2015; **8**: 367.
27. **Patiño-Galindo JÁ, Salvatierra K, González-Candelas F.** Comprehensive screening for naturally occurring hepatitis C virus resistance to direct-acting antivirals in the NS3, NS5A, and NS5B genes in worldwide isolates of viral genotypes 1 to 6. *Antimicrobial Agents and Chemotherapy* 2016; **60**(4): 2402–2416.
28. **European Association for Study of Liver.** EASL recommendations on treatment of hepatitis C 2015. *Journal of Hepatology* 2015; **63**: 199–236.

29. **Lontok E, *et al.*** Hepatitis C virus drug resistance-associated substitutions: state of the art summary. *Hepatology* 2015; **62**(5): 1623–1632.

30. **Tian L, *et al.*** Genotype study on hepatitis C among entry foreigners in China. *Zhonghua Yi Xue Za Zhi* 2015; **95**(12): 903.

31. **Zein NN.** Clinical significance of hepatitis C virus genotypes. *Journal of Hepatology* 1999; **31**(Suppl 1)(2): 65.

32. **Lu L, *et al.*** Hepatitis C virus genotype distribution in China: predominance of closely related subtype 1b isolates and existence of new genotype 6 variants. *Journal of Medical Virology* 2005; **75**: 538–549.

33. **Xia XS, *et al.*** The unique HCV genotype distribution and the discovery of a novel subtype 6u among IDUs co-infected with HIV-1 in Yunnan, China. *Journal of Medical Virology* 2008; **80**: 1142–1152.

34. **Alter MJ.** HCV routes of transmission: what goes around comes around. *Seminars in Liver Disease* 2011; **31**: 340–346.

35. **Nelson PK, Mathers BM, Cowie B, Hagan H.** Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. *The Lancet* 2011; **378**: 571–583.

36. **Awadalla HI, Ragab MH, Nassar NA.** Risk factors of hepatitis C infection among Egyptian blood donors. *Central European Journal of Public Health* 2011; **19**(4): 217–221.

37. **Huang CH, *et al.*** Prevalence and risk factors of hepatitis C among former blood donors in rural China. *International Journal of Infectious Diseases* 2012; **16**: 731–734.

38. **Lin GX, Huang QH, Liu FX.** High prevalence of hepatitis C virus infection in a village in Putian. *Chinese Journal of Experimental and Clinical Infectious Diseases* 2009; **3**(4): 28–31.

39. **Kanwal S, Mahmood T.** Occurrence of genetic modifications in core, 5′UTR and NS5b of HCV associated with viral response to treatment. *Virology Journal* 2014; **11**: 171–191.