The genotype analysis of the hepatitis C virus in Heilongjiang Province, China

Xue-Di Cheng, MD\textsuperscript{a,c}, Hua-Feng Xu, PhD\textsuperscript{b}, Feng Wei, MD\textsuperscript{c}, Li-Xin Jiang, MD\textsuperscript{c}, Hai-Zhou Zhou, PhD\textsuperscript{c,*}

Abstract
Introduction: Hepatitis C virus (HCV) infection is a major public health issue. HCV genotype identification is clinically important to tailor the dosage and duration of treatment, and recombination in intra-patient populations of HCV may lead to the generation of escape mutants, as previously observed for other RNA viruses. Up to now, there is no study assessing HCV genotypes and subtypes in Heilongjiang Province, China.

Methods: To determine genotype and phylogenetic analysis of HCV in Heilongjiang Province is crucial. In this study, we amplified 3 genome regions (5\textsuperscript{′}UTR, E1, and NS5B) of 30 HCV patients in Heilongjiang Province, amplified products were analyzed by bioinformatics.

Results: We found that 23 specimens had concordant subtypes in the 3 gene regions (2a and 1b), 7 HCV patients were considered the recombinants, the recombination pattern of the 7 HCV patients in the 5\textsuperscript{′}UTR, E1, and NS5B region as followed: 1b/2a/1b, 2a/2a/1b, 1b/2a/2a, 1b/2a/1b, 1b/2a/1b, 1b/2a/1b, 2a/2a/1b.

Conclusions: The findings in the present study showed that a higher recombination rate (23\%) than other researches, and the recombination of 2a/1b in the 5\textsuperscript{′}UTR, E1, and NS5B region was only found in the present study up to now.

Abbreviations: 5\textsuperscript{′}UTR = the 5\textsuperscript{′} untranslated region, HCV = hepatitis C virus, PCR = polymerase chain reaction, RT-PCR = reverse transcription polymerase chain reaction.

Keywords: genotype, hepatitis C virus, subtypes

1. Introduction
Hepatitis C virus (HCV) is a blood-borne pathogen. Every year, around 70 to 100 million people are chronically infected with HCV worldwide.\textsuperscript{[1]} Despite the increased availability of therapeutic drugs (IFN+RBV, DAAAs) that have improved the cure rate in recent years, However, drug-resistant variants exist, uncertainty genotyping and recombinations create new problems. Several studies have pointed to the HCV genotype as an important predictive factor of responsiveness to drugs therapy.\textsuperscript{[2]} Different genotypes exhibit different pathogenicity, antigenicity, and response to therapy.\textsuperscript{[3,4]} indicating that the genetic diversity of HCV is involved in diagnosis, pathogenesis, treatment duration, and outcome.\textsuperscript{[5]} Recombination plays a significant role in the evolution of RNA viruses by creating genetic variation,\textsuperscript{[6]} and has the potential to produce “escape mutants”.\textsuperscript{[7]} Hence, determining the genotypes and diversity of HCV would provide valuable insights for diagnosis, prevention, and therapy of HCV infection.\textsuperscript{[3,8]}

There are 7 genotypes and about 100 subtypes of HCV throughout the world.\textsuperscript{[9]} The most commonly detected genotypes in China are 1b and 2a, followed by several less common types.\textsuperscript{[10]} However, the variation in HCV strains circulating in Heilongjiang Province is currently unknown. We analyzed the sequence features of the 5\textsuperscript{′} untranslated region (5\textsuperscript{′}UTR), E1, and NS5B gene regions of HCV isolates circulating in Heilongjiang Province to describe the molecular epidemiology and the genetic diversity of HCV in the area.

2. Materials and methods
2.1. Clinical specimens
Blood samples from 30 HCV cases (female: 12 cases, male: 18 cases) were obtained from the First Affiliated Hospital of Harbin Medical University. The patients all possess major chronic HCV infections, are from 5 different cities of Heilongjiang Province, and transmission of HCV occurred through blood transfusion.
The patients’ ages ranged from 21 to 75 years (median 50 years). All samples were centrifuged at 2000 g for 10 minutes and the serum was collected and stored at −80°C until use.

### 2.2. Viral RNA extraction and nested reverse transcription polymerase chain reaction (RT-PCR) amplification

Viral RNA was extracted from serum using the Viral RNA Mini Spin Kit (Qiagen, Hilden, Germany). Reverse-transcribed PCR was performed using random hexanucleotide primers. For the amplification of HCV cDNA, nested polymerase chain reaction (PCR) assays were used for the individual amplifications of the 5′UTR (225 bp, nt 85–310, GenBank accession number AB249644), E1 (473 bp, nt 843–1315, GenBank accession number AF009606), and NS5B (339 bp, nt 8002–8340, GenBank accession number D00944). Nested PCR was performed per the following: initial denaturation at 94°C for 5 minutes followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for NS5B, 58°C for E1 for 35 seconds, and extension at 72°C for 1 minute, followed by a final extension step at 72°C for 10 minutes. For the second round of the PCR, the same cycle programs were performed using internal sense and antisense primers of the proper region. However, the annealing at NS5B and E1 regions were performed at 58°C and 62°C. For the 5′UTR region, all conditions were the same as for NS5B and E1, except the annealing temperature was at 56°C. The PCR products of 5′UTR, E1, and NS5B were purified and sequenced directly. All primer sequences are shown in Table 1.

### 2.3. Sequencing and phylogenetic analysis

The sequences of 5′UTR, E1, and NS5B were compiled, aligned, and adjusted to the reference sequences from an HCV database (https://hcv.lanl.gov/) using the CLUSTAL X program. Neighbor-joining trees were made using MEGA version 6.0. The reliability was estimated by 1000 bootstrap replications.

### 2.4. Recombination analysis

Putative recombinant sequences were identified with the SimPlot program, using concatenated (5′UTR, E1, and NS5B) sequences and the reference sequences obtained from the HCV sequence database. This program is based on a sliding window method and constitutes a way of graphically displaying the coherence of the sequence relationships over the entire length of a set of aligned homologous sequences. The window width and the step size were set to 100 bp and 10 bp, respectively.

### 3. Results

#### 3.1. Genotyping of HCV isolates from Heilongjiang Province

In this study, HCV genotyping was carried out for 30 patients from different districts of Heilongjiang by using 5′UTR, E1, and NS5B sequences. The sequences were aligned using the CLUSTAL X program. We conducted a phylogenetic analysis using the MEGA software tool (Fig. 1). As a measure of the robustness of each node, we used the bootstrap method (1000 pseudo-replicates). Twenty-three patients had consistent results in 3 regions: 15 were subtype 2a, 8 were subtype 1b, and 7 patients were considered recombinants (Fig. 1). The results are summarized in Table 2.

### 3.2. Recombination analysis

Conflicting phylogenetic data, such as that obtained for the 5′UTR, E1, and NS5B sequences in samples HLJHCV9, HLJHCV21, HLJHCV23, HLJHCV26, HLJHCV28, HLJHCV34, and HLJHCV62 are indicative of the presence of either a mixed infection with 2 different genotypes or an intergenotypic recombinant virus (Fig. 1). In the phylogenetic analysis of the 5′UTR and E1 regions, samples HLJHCV21 and HLJHCV62 were subtype 2a, whereas these samples in the NS5B region were subtype 1b. The sequences of HLJHCV9, HLJHCV26, HLJHCV28, and HLJHCV34 from 5′UTR and NS5B regions belonged to subtype 1b. However, the same samples in the E1 region were rather similar to the sequences of subtype 2a. In the phylogenetic analysis of the E1 and NS5B regions, sample HLJHCV23 belonged to subtype 2a, and belonged to subtype 1b in the 5′UTR region. The similarity analysis was performed using the SimPlot application, which gave the corresponding region which each subtype belonged to. The recombination pattern of the 7 HCV patients (HLJHCV9, HLJHCV21, HLJHCV23, HLJHCV26, HLJHCV28, HLJHCV34, and HLJHCV62) in the 5′UTR, E1, and NS5B regions was as follows: 1b/2a/1b, 2a/2a/2a, 2a/2a/1b, 1b/2a/1b, 1b/2a/1b, 1b/2a/1b, and 2a/2a/1b.

To identify a possible recombinant event, we also analyzed the break and joining points of the 7 recombinants using the SimPlot software. The break points concentrated mainly at the locations 110 to 125 and 575 to 598 in the 3 HCV regions (5′UTR, E1, and NS5B) (Table 3).

### 4. Discussion

HCV is an RNA virus with a high rate of genetic mutation. HCV genotyping serves as a predictor of antiviral treatment duration and also response to therapy. The distribution of the HCV genotypes differs geographically. Genotypes 1a, 1b, 2a, 2b,
and 3a are commonly seen throughout the world, whereas genotypes 4, 5, and 6 are only detected in certain geographic regions. Subtypes 3a and 1b are the most commonly found in China, followed by subtypes 3b and 2a in Ningxia. The most prevalent subtype in Wuhan is 1b (71.98%), followed by...
genotypes 2, 3, and 6. In Hubei, the predominant genotypes are 1b and 2a. Until this study, little was known about the distribution of HCV genotypes in Heilongjiang Province. Our results are consistent with the distribution of HCV genotypes in the north, where subtypes 1b and 2a are the main circulating subtypes. Our study demonstrated 23% (7/30) recombination, which contradicts some reports indicating that recombination is a rare phenomenon. As the probability of detecting recombinant strains is low, it is likely that their incidence is underestimated.

Guidelines have been proposed for the classification of genotypes/subtypes, using either the entire genome or the core/E1 and NS5B regions of HCV. 5'UTR is the region of choice for qualitative and quantitative HCV RNA detection due to its high level of conservation and sensitivity. For this reason, it has most often been used by clinical laboratories for routine genotyping of HCV. However, due to this high level of conservation, the 5'UTR is limited in its ability to discriminate genotype 6 from genotype 1 and subtypes within genotypes 1, 2, 3, 4, and 6. The 5'UTR region also has several polymorphisms that may frequently prevent HCV subtyping or result in misclassification of a significant percentage of subtype 1a isolates as subtype 1b. The E1 region is occasionally used because it is hypervariable. Our previous research demonstrates that the E1 protein may play an important role in the immunogenicity of HCV, especially in HCV subtype 1b. Genotyping based on sequencing ≥2 HCV genomic regions (C/E1, NS5A, NS5B), exhibited excellent concordance in another study. The NS5B region is easier to subtype, which makes up for the defect in the 5'UTR region. Subtyping using a single segment is therefore not reliable and using multiple segments is advisable. We employed the 5'UTR, E1, and NS5B regions in subtyping analysis, allowing us to detect recombination.

HCV, similar to other RNA viruses, exploits all possible mechanisms of genetic variation to ensure its survival. Recombination plays a significant role in the evolution of RNA viruses by creating genetic variation. Both inter- and intra-genotypic recombination have been reported in HCV populations in different geographic locations (Table 4), and the rate of recombination ranges from 0.8% to 5%. However, some studies, including ours, report a higher rate (10%–56%). We sequenced the PCR product directly in our study, the sample sequencing peaks are single and uniform, this excludes mixed
infections. Despite the growing evidence for recombination in RNA viruses, the evolutionary reasons for its occurrence remain uncertain. Our study showed recombination in males (6/18, 33%), which was more than 4-fold higher than that in females (1/12, 8%). A higher HCV prevalence in elderly males has been reported in other studies. Similarly, our preliminary study shows that older men have a higher rate of HCV infection than women (data not shown). Indeed, recombination has been associated with the expansion of viral host range, increases in virulence, the evasion of host immunity, and the evolution of resistance to antivirals. Susceptibility to recombination may be the cause of the higher rate in males as compared with females.

HCV genotypes are considered predictors of the outcome of interferon treatment. Interferon resistance has been attributed to sequence variability of the NS5A protein (amino acid residues 2209–2248) and of the E2 region (5, 18, 31). Variation is the main cause of the recombination, and the break points are mainly the locations of the variation and recombination. However, the break points concentrated mainly at locations 110 to 125 and 575 to 598 in the 3 HCV regions (5'UTR, E1, and NS5B) in our study. Our study demonstrates that analysis of more than one sub-genomic region is necessary to avoid missing recombinant strains, and it suggests that we should further study these positions.

Two out of the 7 recombinant cases in our study were subtype 2a/1b, which has not previously been reported. Genotype 1 is more resistant to IFN therapy than genotypes 2 and 3. The antiviral response to treatment for HCV inter-genotypic recombinant genotype 2/1 has been reported to be similar to responses for genotype 1 patients as compared with genotype 2 patients. The new recombinant pattern may be a greater challenge to HCV treatment. When treatment response is not as expected, recombination and existence of an inter-genotypic recombinant strain could be screened for as a source of the response.

5. Conclusions

Consistent with the most commonly detected genotypes in China, HCV subtypes 2a and 1b are the main subtypes in Heilongjiang Province. In this area, there is a higher recombination rate than in other studies, and the 2a/1b recombinant in the 5'UTR, E1, and NS5B regions is reported for the first time. This recombinant pattern may prove more difficult to treat. To obtain more accurate results, we recommend ≥2 HCV genomic regions be adopted for classification.

Acknowledgments

We thank the participants and staff of the Department of Laboratory Diagnosis of the First Affiliated Hospital of Harbin Medical University, without whose generosity none of this work would have been possible.

Author contributions

Conceptualization: Xue-Di Cheng, Hai-Zhou Zhou.
Data curation: Xue-Di Cheng.
Formal analysis: Xue-Di Cheng.
Investigation: Xue-Di Cheng, Hua-Feng Xu, Hai-Zhou Zhou.
Methodology: Xue-Di Cheng.
Software: Li-Xin Jiang.
Writing – original draft: Feng Wei.

References

[1] Asselah T, Hassanein T, Waked I, et al. Eliminating hepatitis C within low income countries—the need to cure Genotypes 4, 5, 6. J Hepatol 2017;11:37.
[2] Broor S, Bhupatiraju AV, Broor SL, et al. Analysis of genetic variability of Indian isolates of hepatitis C virus. Arch Virol 2004;149:1185–92.
[3] Zein NN. Clinical significance of hepatitis C virus genotypes. Clin Microbiol Rev 2000;13:223–35.
[4] Legrand-Abravanel F, Colson P, Leguillou-Guillemette H, et al. Influence of the HCV subtype on the virological response to pegylated interferon and ribavirin therapy. J Med Virol 2009;81:2029–35.
[5] Bolcic F, Lafer N, Trinchero J, et al. A clustering phenomenon among HCV-1a strains among patients coinfected with HIV from Buenos Aires, Argentina. J Med Virol 2012;84:570–81.
[6] Colina R, Casane D, Vasquez S, et al. Evidence of intratypic recombination in natural populations of hepatitis C virus. J Gen Virol 2004;85:31–7.
[7] Georgescu MM, Delpeyroux F, Tardy-Panit M, et al. High diversity of poliovirus strains isolated from the central nervous system from patients with vaccine-associated paralytic poliomyelitis. J Virol 1994;68:8089–101.
[8] Qi Y, Chen Q, Hao F, et al. Subtype distribution of hepatitis C virus in Fujian, China. J Med Virol 2015;88:498–505.
[9] Laidlaw SM, Dustin LB. Interferon lambda: opportunities, risks, and uncertainties in the fight against HCV. Front Immunol 2014;5:545.
[10] Tian D, Li L, Liu Y, et al. Different HCV genotype distributions of HIV-infected individuals in Henan and Guangxi, China. Plos One 2012;7:e50343.
[11] Murphy DG, Willems B, Deschenes M, et al. Use of sequence analysis of the NS5B region for routine genotyping of hepatitis C virus with reference to C/E1 and 5 untranslated region sequences. J Clin Microbiol 2007;45:1102–12.
[12] Cantaloube JF, Gallian P, Attoui H, et al. Genotype distribution and molecular epidemiology of hepatitis C virus in blood donors from southeast France. J Clin Microbiol 2005;43:3624–9.
[13] Gededzha MP, Muzeze M, Burnett RJ, et al. Complete genome analysis of hepatitis B virus in human immunodeficiency virus infected and uninfected South Africans. J Med Virol 2016;88:1560–6.

[14] Lole KS, Bollinger RC, Paranjape RS, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 1999;73:152–60.

[15] Idrees M, Butt S, Awan Z, et al. Nucleotide identity and variability among different Pakistani hepatitis C virus isolates. Virol J 2009;6:130.

[16] Triantos C, Konstantakis C, Tselekouni P, et al. Epidemiology of hepatitis C in Greece. World J Gastroenterol 2016;22:8094–102.

[17] Wu Z, Cui L, Zhao W, et al. Molecular epidemiology of hepatitis C infections in Ningxia, China: genotype, phylogeny and mutation analysis. Virol J 2016;13:172.

[18] Niu Z, Zhang P, Tong Y. Age and gender distribution of hepatitis C virus prevalence and genotypes of individuals of physical examination in Wuhan, Central China. Springerplus 2016;5:1557.

[19] Peng J, Lu Y, Liu W, et al. Genotype distribution and molecular epidemiology of hepatitis C virus in Hubei, Central China. Plos One 2015;10:e0137059.

[20] Viazov S, Widell A, Nordenfelt E. Mixed infection with two types of hepatitis C virus in human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 1999;73:152–60.

[21] Simmonds P, Bukh J, Combet C, et al. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology 2005;42:962–73.

[22] Chen Z, Weck KE. Hepatitis C virus genotyping: interrogation of the 5′ untranslated region cannot accurately distinguish genotypes 1a and 1b. J Clin Microbiol 2002;40:3127–34.

[23] Avo AP, Agua-Doce I, Andrade A, et al. Hepatitis C virus subtyping based on sequencing of the C/E1 and NS5B genomic regions in comparison to a commercially available line probe assay. J Med Virol 2013;85:815–22.

[24] Cheng XD, Xu HF, Wei XM, et al. Variation analysis of E1 and E2 in hepatitis C virus subtypes. Arch Virol 2015;160:2479–256.

[25] McCormick AL, Macartney MJ, Abdi-Abshir I, et al. Evaluation of sequencing of HCV core/E1, NS5A and NS5B as a genotype predictive tool in comparison with commercial assays targeting 5′ untranslated region cannot accurately distinguish genotypes 1a and 1b. J Clin Microbiol 2002;40:3127–34.

[26] Cristina J, Colina R. Evidence of structural genomic region recombination in hepatitis C virus. Virol J 2006;3:53.

[27] Gonzalez-Candelas F, Lopez-Labrador FX, Bracho MA. Recombination in hepatitis C virus. Viruses 2011;3:2006–24.

[28] Simmonds P, Bukh J, Combet C, et al. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology 2005;42:962–73.

[29] Chen Z, Weck KE. Hepatitis C virus genotyping: interrogation of the 5′ untranslated region cannot accurately distinguish genotypes 1a and 1b. J Clin Microbiol 2002;40:3127–34.

[30] Avo AP, Agua-Doce I, Andrade A, et al. Hepatitis C virus subtyping based on sequencing of the C/E1 and NS5B genomic regions in comparison to a commercially available line probe assay. J Med Virol 2013;85:815–22.

[31] Cheng XD, Xu HF, Wei XM, et al. Variation analysis of E1 and E2 in hepatitis C virus subtypes. Arch Virol 2015;160:2479–256.

[32] Gaspareto KV, Ribeiro RM, de Mello Malta F, et al. HCV inter-subtype 1a/1b recombinant detected by complete-genome next-generation sequencing. Arch Virol 2016;161:2161–8.

[33] Lee YM, Lin HJ, Chen YJ, et al. Molecular epidemiology of HCV genotypes among injection drug users in Taiwan: full-length sequences of two new subtype 6w strains and a recombinant form_2b6w. J Med Virol 2010;82:57–68.

[34] Legrand-Abravanel F, Claudignon J, Nicot F, et al. New natural intergenotypic (2/5) recombinant of hepatitis C virus. J Virol 2007;81:4357–62.

[35] Moreno P, Alvarez M, Lopez L, et al. Evidence of recombination in hepatitis C virus populations infecting a hemophilic patient. Virol J 2009;6:203.

[36] Kageyama S, Agdamag DM, Alesna ET, et al. A natural inter-genotypic (2b/1b) recombinant of hepatitis C virus in the Philippines. J Med Virol 2006;78:1423–8.

[37] Sentandreu V, Jimenez-Hernandez N, Torres-Puente M, et al. Evidence of recombination in intrapatient populations of hepatitis C virus. Plos One 2008;3:e3239.

[38] Chen YD, Liu MY, Yu WL, et al. Mix-infections with different genotypes of HCV and with HCV plus other hepatitis viruses in patients with hepatitis C in China. World J Gastroenterol 2003;9:984–92.

[39] Takahashi H, Mizuta T, Eguchi Y, et al. Post-challenge hyperglycemia is a significant risk factor for the development of hepatocellular carcinoma in patients with chronic hepatitis C. J Gastroenterol 2011;46:790–8.

[40] Xu H, Yu G, Sun H, et al. Use of parenteral caffeinum natrio-benzoicum: an underestimated risk factor for HCV transmission in China. BMC Public Health 2015;15:928.

[41] Brown DW. Threat to humans from virus infections of non-human primates. Rev Med Virol 1997;7:239–46.

[42] Gibbs MJ, Weiller GF. Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. Proc Natl Acad Sci USA 1999;96:8022–7.

[43] Khatchikian D, Orlich M, Rott R. Increased viral pathogenicity after insertion of a 285 ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. Nature 1989;340:156–7.

[44] Malim MH, Emerman M. HIV-1 sequence variation: drift, shift, and attenuation. Cell 2001;104:469–72.

[45] Nora T, Charpentier C, Tenaillon O, et al. Contribution of recombination to the evolution of human immunodeficiency viruses expressing resistance to antiretroviral treatment. J Virol 2007;81:7620–8.

[46] Taylor DR, Shi ST, Lai MM. Hepatitis C virus and interferon resistance. Microbes Infect 2000;2:1743–56.

[47] Vierling JM. Hepatitis C virus viral assays in the direct-acting antiviral era. Clin Liver Dis 2013;17:27–45.

[48] Hedskog C, Doehle B, Chodavarapu K, et al. Characterization of hepatitis C virus intergenotypic recombinant strains and associated virological response to sofosbuvir/ribavirin. Hepatology 2015;61:471–80.

[49] Morel V, Fournier C, Francois C, et al. Genetic recombination of the hepatitis C virus: clinical implications. J Viral Hepat 2011;18:77–83.

[50] Calado RA, Rocha MR, Parreira R, et al. Hepatitis C virus subtypes circulating among intravenous drug users in Lisbon, Portugal. J Med Virol 2011;83:608–15.