

LETTER TO THE EDITOR

Distribution of HCV Genotypes in Tianjin, China and Characterization of a Full-Length Viral Genome Representing Subtype 6xa

Wei Hou

Tianjin Second People's Hospital and Tianjin Institute of Hepatology, Tianjin, China

(Clin. Lab. 2019;65:xx-xx. DOI: 10.7754/Clin.Lab.2019.190150)

Correspondence:

Wei Hou
Tianjin Second People's Hospital and
Tianjin Institute of Hepatology
7 Sudi South Road
Nankai District, Tianjin
China
Phone: +86 22 27428929
Email: houweicn@163.com

KEY WORDS

hepatitis C virus, genotype, subtype 6xa

LETTER TO THE EDITOR

Hepatitis C virus (HCV) is a major cause of chronic hepatitis C (CHC) that can lead to liver cirrhosis and hepatocellular carcinoma (HCC). An estimated 130 to 170 million people are chronically infected by HCV worldwide and China has the most people with HCV (29.8 million) [1]. Seven major genotypes (GTs) and multiple subtypes have been recognized to date [2-4]. HCV GTs are associated with a marked divergence in geographical distribution in mainland China [5-8]. However, the distribution of viral genotypes and subtypes in patients with HCV infection in Tianjin, the biggest coastal city in north China with the population more than 15 million, is still largely unknown.

To investigate genetic diversity and distribution of HCV in Tianjin, 1,228 chronic HCV patients from Tianjin Second People's Hospital and Tianjin Institute of Hepatology between 2013 and 2016 were enrolled in this cross-sectional study. HCV genotypes and subtypes were determined by phylogenetic analyses based on the sequences of the nonstructural protein 5B (NS5B) genomic regions. Five HCV genotypes (1, 2, 3, 4, and 6) with seven subtypes (1a, 1b, 3a, 3b, 4a, 6a, and 6xa) were identified (Figure 1A). HCV genotype 1b (869 cases, 70.8%) was the most prevalent genotype, followed by 2a (256 cases, 20.8%), 3a (72 cases, 5.8%), 3b (21 cases, 1.7%), 6a (5 cases, 0.4%), 1a (2 cases, 0.2%), 4a (2 cases, 0.2%), 6xa (1 case, 0.1%) (Figure 1B). To better understand genetic variation and evolution of the uncommon isolate 6xa identified in Tianjin, full-

Letter to the Editor accepted February 2, 2019

Table 1. The number of nucleotides in genomic regions of TJ6xa corresponding to the reference isolate H77 (GenBank accession no. AF009606). Protein translation of TJ6xa is highlighted.

CDS	Nucleotide position relative to CDS start in reference isolate H77	Nucleotide position relative to TJ6xa sequence start
5UTR	35 → 341	1 → 307
Core	342 → 914 (191aa)	308 → 880 (191aa)
MSTLPKPQRLTKRNTNRRPQNVKFPGGGQIVGGVYLLPRRGPKLGVAVRKTSERSQPRGRRQPIPKQRPPTGRHWAQPGYPWPLY GNEGCGWAGWLLSPRGRSPHWGPNDRRRSRNLGKVIDTLTCGFADLMGYIPVIGAPLGGVATALAHGVRAVEDGINYATGNLPGC SFSIFLLALLSCLTVPASA		
E1	915 → 1490 (192aa)	881 → 1456 (192aa)
VHYSNKSGLYLTDNCPNASIVYQAHDVILHLPGCVPCVKAGNQSRCWYPASPTLAIIPNVSVVPRGFRRHVDLLVGAFAFCSSMYVG DLCGGLFLIGQLFTFAPQHHQITQECNCISIYTGHSVGRHMAWDMMMNWSPTVGYVVSLLRVPQILMDIVLEGHWGVXGALLYFS MVANWAKVAAVLILFAGVDA		
E2	1491 → 2579 (363aa)	1457 → 2566 (370aa)
ETHXTGAAAGRAAXGFAGLFSAGSRQNLQLINTNGSWHVNRTALNCNDSLQTGFLAGLFYRHKFNSSGCPERLSSCKSLTYFDQGW GKLTIANVTGSSQDRPYCWHYAPNSCNVIPAAASVCGPVYCFPTSPVVVGTDRKGFPTHTWGENESDVFLGSMRPPAGGWYGCCTW MNSTGFTKTCGAPPCQIRXPINSSXTNTLLCPTDCFRKHPDATYKCKGSGPWLTTPRCLVXYPYRLWHYPTVNFVHKVRLYVGG VEHRFDAACNWTRGERCELDDRDRIEMSPLLFSTTQLSILPCSFTTTPALSTGLIHLHQNVVDVQYLYGVSSNIVSWAVRWEYLVLAF LVLADARVCAVFWLMMVAQAEA		
p7	2580 → 2768 (63aa)	2567 → 2755 (63aa)
ALENLILNAASAASNQGWVWCLLXVCCAWYIKGRAVPAFTYGLLQLWPLLLLLLALPPRAFA		
NS2	2769 → 3419 (217aa)	2756 → 3406 (217aa)
WDGEQAAXMGALVTXGVTIFLSPXYKSMILHLIWWLQYFIARAEAIQVWVPSLHVRGGRDGGXIVLTTLLHPSLGFVTKILLAILG PLYLLQGALVRVPYFVRAHALLRVCGLVRGVLGGKYIQAALMIKLGALGTGTIYDHLSPSDWAASGLRDLAVAVEPVVFPMEKKVI TWGADTAACGDILHGLPVSARKGSLVLLGPADDMQGRGWKLL		
NS3	3420 → 5312 (631aa)	3407 → 5299 (631aa)
APVTAYAQQTRGVLTIVTSLTGRDKNEVEGEVQVSTATQTFLATAINGVLWTVFHGAGSKTLAGPKGPVCOMYTNVDQDLVGV PAPPGCKSLTPTCCGASDLYLVTRNADVIPARRRGDTRASLLSPRPLATLKGSSGGPVLCPSGHIVGLFRAAVCTRGVAKALDFVAVEN METTMRSPVFTDNSSPPAVPSTYQVGYLHAPTGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGFAGYMSRAHGIEANIRTGVRTITTG GAIYTYGKFLADGGCSGGAYDIIICDECHSTDPTTVLIGITVLDQAEATAGVRLTVLATATPPGSVTVPHPNITETALPTTGEVVPYGR AIPLDFIKGGRHLIFCHSKKKCDELSSQLRSLGINAVAFYRGLDVSVIPTSGLDGVVVCATDALMTGYTGDFDSVIDCNVAVTQTVDFSLD PTFSIETTVPQDAVSRSQRRGRTGRGKPGIYRFVSKGERPSGMFDTVVLAAYDTGLAWYELTPSETTVRLRAYLNTPLPVCQDHL EFWEGVFTGLTHIDAHFLSQTKQGGENFAYLVAQATVCARAKAPPCWDVMWKCLSRLKPMVLGPTPLLYRLGPVQNDIHTHPIT KYIMTCMSADLEVIT		
NS4A	5313 → 5474 (54aa)	5300 → 5461 (54aa)
STWVLVGGVLAALAAAYCLTVGCVVICGRIVLSGKPAVVPDREVLYQQFDEMEEC		
NS4B	5475 → 6257 (261aa)	5462 → 6244 (261aa)
SRHLPYLAEGQQLAEQFKQKVLGLLQSTAKQAEELKPAVHATWPKLEQFWQKHMWNFVSGIQYLAGLSTLPGNPAIASLMAFSASL TSPLSTSTLLNILGGWVASQLATPAGSTSFVVCGLAGAAVGSVGLGKVLIDILAGYGAGVSGALVCFKIMSGETPTMEDMANLLPA LLSPGALVGVVCAAILRRHVGPSEGAQWMNRLIAFASRGNHVSPTHYVPETDASRQVISILSSLTITSLRRLHQWISEDCSMP		
NS5A	6258 → 7601 (448aa)	6245 → 7600 (452aa)
STSWLRDIWDWVCTVLSDFKTWLKAKLVPSLPGVPFISQCRGVRGVWKGGINHTTCCPGANITGHVKNKSMRVLGPRTCNSNVWQ GTFPINATTTGPSVIPAPNYSRALWRIAAEEYAEVRXVGDYHYIVGVTNDNLKCPQVPSPEFFTEVDGVRHLHRYAPACKPLLRDEV FSVGLNSYVVGSQLPCEPEPDVAVLTSMLTDPHITAETAAGRLLARGSPSLASSXASQLSAPSLKATCTARHDSPADLIEANLLWRQ EMGNGNITRVESENKVVVLDSDPLVPEYDTRPSVPAECHRRPPKFPALPIWARPDYNPPLLENWKRPEYSAPVHGCMPMPRET PIPPRRKRLVRLDESTVAALSELAERSFPTTSTGSDKTADSGVPTTGTIPPTPPGDDEAASDAGSYSSMPLEGEPPDLSSGSWSTV SEEDDSVVCC		
NS5B	7602 → 9377 (591aa)	7601 → 9376 (591aa)
SYSYSWTGALITACAAEEKLPISPLSNLIRHNLVYSTTSRSAALRQKKVTFDRQLQVVDQYYYDTLKEIKLRASKVSAAGLLSVEEAC DLTPPHSARSQFGYGAQDVRSHASKAVNHINSVWDLLEDKTTPIPTTIMAKNEVFCVEPSKGGXKPARLIVYDGLGVRVCEKMALY DITRKLPTAVMGAAAYGFQYSPKDRVDHLLKLRSKKTPMGFSYDTRCFDSTVTERDIRTEADYLSQQLQPEARDAITSLTERLYCGG PMFNKQSCGYRRRCRASGVLPTSMGNTITCYLKAQAACKAAGLKDFFDMLVCGDDLVVICESAGVQEDVAALRAFTEAMTRYAPP GDDPHPEYDLELITSCSSNVSAIDDTGKRYYYLTRDPTIPFARAAWETARHTPVNSWLGNIIMYAPTIWVRMVMTHFFLILQAQEV LGNPLDFDMYGVTVSPTPLDPAHQRLHGMAAFSLHNSPGLNVRVGAACLRKLGAPPLRAWHRARAVRAKLIQAQGGKAAICGKY LFWAVRTRKSKLTPLRDASMLDLGWFSTGFSGGDIYHSVSRARPRLLLLCLLLLTVGVGIFLLPARX		
3UTR	9378 → 9646	9377 → 9413

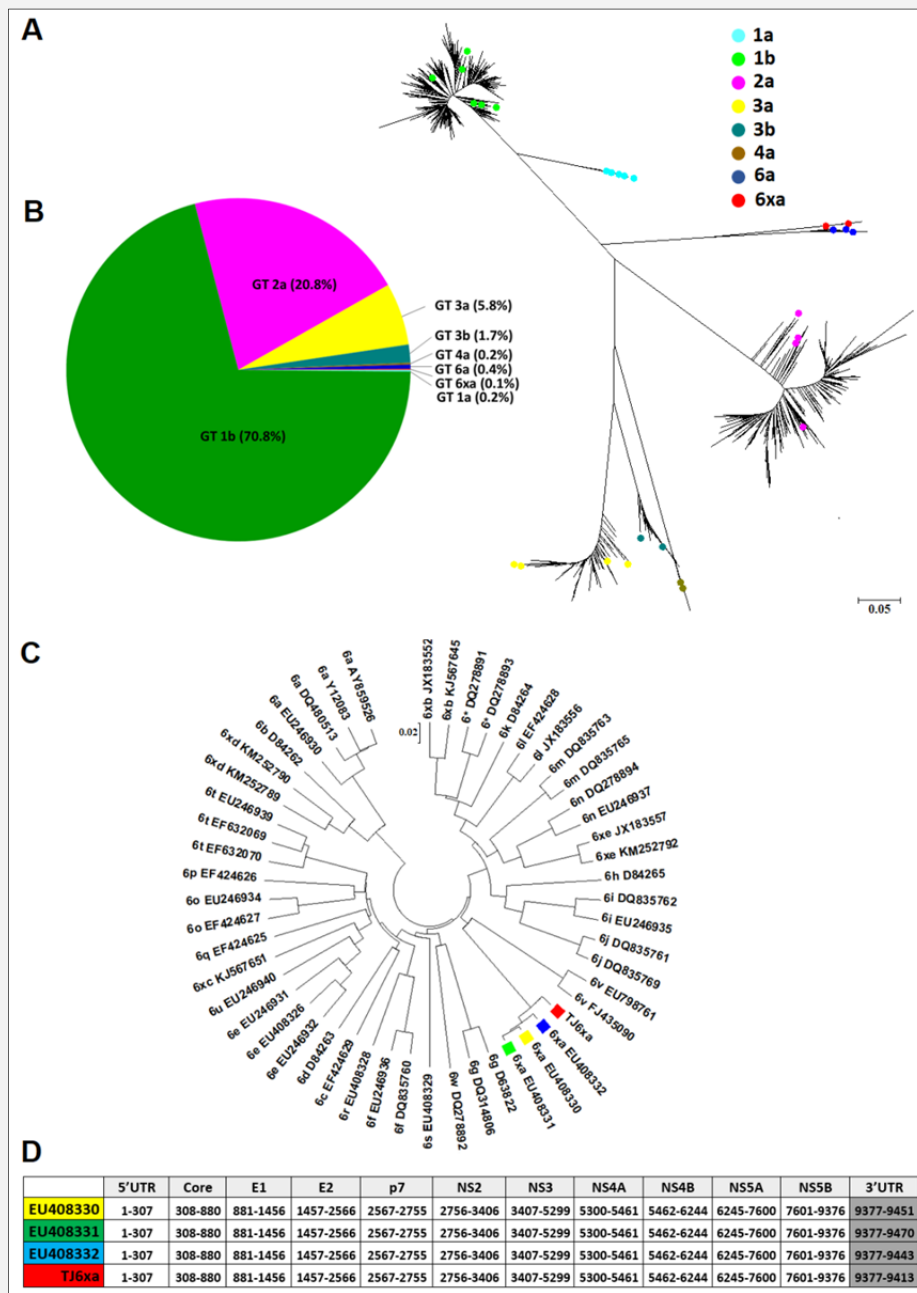


Figure 1. Distribution of HCV genotypes and subtypes in Tianjin and genomic characterization of the isolate TJ6xa.

(A) Phylogenetic tree is constructed by the neighbor joining method using MEGA6, based on the NS5B sequences from 1228 subjects. Genotype/subtype of reference sequences (Subtype 1a: AF511950, EF407419, EF407457, M62321, NC_004102; Subtype 1b: D11355, D90208, EF032892, EU781827, EU781828, M58335; Subtype 2a: AB047639, AY746460, D00944, HQ639944; Subtype 3a: AF046866, D17763, D28917, JN714194; Subtype 3b: D49374, JQ065709; Subtype 4a: DQ418789, DQ988074; Subtype 6a: AY859526, DQ480513, EU246930; Subtype 6b: EU408331, EU408332) indicated with different colored dots, is shown in each branch. (B) Schematic presentation of the genotype/subtype distribution in Tianjin. (C) Phylogenetic tree is constructed by the neighbor joining method using MEGA6, based on full-length genomic sequences of TJ6xa and fifty-one genotype 6 references (Subtype 6a: Y12083, AY859526, DQ480513, EU246930; Subtype 6b: D84262; Subtype 6c: EF424629; Subtype 6d: D84263; Subtype 6e: EU246931, EU246932, EU408326; Subtype 6f: DQ835760, EU246936; Subtype 6g: D63822, DQ314806; Subtype 6h: D84265; Subtype 6i: DQ835762, EU246935; Subtype 6j: DQ835761, DQ835769; Subtype 6k: D84264; Subtype 6l: EF424628, JX183556; Subtype 6m: DQ835763, DQ835765; Subtype 6n: DQ278894, EU246937; Subtype 6o: EF424627, EU246934; Subtype 6p: EF424626; Subtype 6q: EF424625; Subtype 6r: EU408328; Subtype 6s: EU408329; Subtype 6t: EF632069, EF632070, EU246939; Subtype 6u: EU246940; Subtype 6v: EU798761, FJ435090; Subtype 6w: DQ278892; Subtype 6x: EU408330, EU408331, EU408332; Subtype 6y: JX183552, KJ567645; Subtype 6zc: KJ567651; Subtype 6xd: KM252789, KM252790; Subtype 6xe: JX183557, KM252792; Subtype 6*: DQ27889, DQ278893). (D) The number of nucleotides in each genomic region of HCV 6xa isolates.

length genomic sequences of the isolate designated TJ6xa were successfully obtained by deep sequencing. A phylogenetic tree was constructed using full-length genomic sequences of TJ6xa and fifty-one genotype 6 references (Figure 1C). The full size genome of TJ6xa with 9,413 nt contained a single ORF (308 - 9,376) with ten protein-coding regions as follows: core (308 - 880), E1 (881 - 1,456), E2 (1,457 - 2,566), p7 (2,567 - 2,755), NS2 (2,756 - 3,406), NS3 (3,407 - 5,299), NS4A (5,300 - 5,461), NS4B (5,462 - 6,244), NS5A (6,245 - 7,600), and NS5B (7,601 - 7,396) (Figure 1D, Table 1). Compared with three HCV 6xa reference isolates (EU-408330, EU408331, and EU408332), only the 3'UTR was variable in length (Figure 1D).

Overall, to the best of our knowledge, this study is the largest yet to describe the distribution of HCV genotypes and subtypes in Tianjin, the biggest coastal city in north China. The predominance of subtype 1b, followed by 2a in Tianjin, is in concordance with the previous studies performed in different regions and different populations of China [5-8]. Our results provide further evidence to indicate that the influx of uncommon isolates such as 4a and 6xa has been modifying the epidemiological pattern in northern China, since genotype 4 is mainly found in the Middle East and North Africa, while genotype 6 is primarily distributed in southern China and Southeast Asia [2-4]. These findings provided new insight into the geographical distribution of HCV variants, which will support individualized programs of surveillance and prevention of HCV infection in this area.

Acknowledgment:

WH was supported by the National Natural Science Foundation of China (grant nos. 30800974 and 81271845) and Tianjin Municipal Health Bureau of Science and Technology Fund (grant nos. 2012KR02 and 12KG118).

Declaration of Interest:

There are no conflicts of interest.

References:

- Hajarizadeh B, Grebely J, Dore GJ. Epidemiology and natural history of HCV infection. *Nat Rev Gastroenterol Hepatol* 2013; 10:553-62 (PMID: 23817321).
- Smith DB, Bukh J, Kuiken C, et al. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology* 2014;59: 318-27 (PMID: 24115039).
- Bukh J. The history of hepatitis C virus (HCV): Basic research reveals unique features in phylogeny, evolution and the viral life cycle with new perspectives for epidemic control. *J Hepatol* 2016;65:S2-S21 (PMID: 27641985).
- Welzel TM, Bhardwaj N, Hedskog C, et al. Global epidemiology of HCV subtypes and resistance-associated substitutions evaluated by sequencing-based subtype analyses. *J Hepatol* 2017;67: 224-36 (PMID: 28343981).
- Du G, Li X, Musa TH, et al. The nationwide distribution and trends of hepatitis C virus genotypes in mainland China. *J Med Virol* 2019;91:401-10 (PMID: 30192393).
- Zhang Y, Chen LM, He M. Hepatitis C Virus in mainland China with an emphasis on genotype and subtype distribution. *Virology* 2017;14:41 (PMID: 28231805).
- Chen Y, Yu C, Yin X, Guo X, Wu S, Hou J. Hepatitis C virus genotypes and subtypes circulating in Mainland China. *Emerg Microbes Infect* 2017;6:e95 (PMID: 29089588).
- Huang K, Chen J, Xu R, et al. Molecular evolution of hepatitis C virus in China: A nationwide study. *Virology* 2018;516:210-8 (PMID: 29407379).