# Genome Sequence of the Hepatitis C Virus Subtype 6n Isolated from Malaysia

Kim Tien Ng,[a] Yeat Mei Lee,[a] Haider Abdulrazzaq Abed Al-Darraji,[a] Xueshan Xia,[b] Yutaka Takebe,[a,c] Kok Gan Chan,[d] Ling Lu,[e] Sanjiv Mahadeva,[f] Adeeba Kamarulzaman,[a] Kok Keng Tee[a]

Centre of Excellence for Research in AIDS (CERiA), Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia[a]; Faculty of Environmental Science and Engineering and Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming, Yunnan, China[b]; AIDS Research Center, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo, Japan[c]; Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia[d]; Viral Oncology Center, Department of Pathology and Laboratory Medicine, University of Kansas, Medical Center, Kansas City, Kansas, USA[e]; Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia[f]

K.T.N. and Y.M.L. contributed equally to this article.

We report the full genome sequence of hepatitis C virus (HCV) subtype 6n from Kuala Lumpur, Malaysia. Phylogenetic analysis of the isolate 10MYKJ032 suggests that Southeast Asia might be the origin for the HCV subtype 6n and highlights the possible spread of this lineage from Southeast Asia to other regions.

Address correspondence to Kok Keng Tee, k2tee@um.edu.my.

Hepatitis C virus (HCV), a positive-sense RNA virus from genus *Hepacivirus* of the family *Flaviviridae*, is a predominant and widely distributed bloodborne pathogen associated with liver cirrhosis and hepatocellular carcinoma in humans (1). According to the World Health Organization, HCV infects approximately 150 million people worldwide (2), of which 10 million are injecting drug users (3). To date, there are six confirmed genotypes, denoted genotypes 1 to 6 (4), and a provisionally assigned genotype 7 (5), with each genotype diverging further into multiple subtypes with distinct genetic variability. Such genetic diversity has been significantly associated with geographical and epidemiological factors (4, 6, 7). Genotypes 1, 2, and 3 are distributed globally; genotype 4 is confined to the Middle East and Central Africa; genotype 5 is predominantly found in South Africa; and genotype 6, the most divergent genotype, is largely found in Southeast Asia (8, 9). Here, we describe the first genome sequence of HCV subtype 6n isolated in Kuala Lumpur, Malaysia, from a consenting 37-year-old HIV-1-infected homosexual male with a history of injecting drug use.

A plasma sample was collected in July 2010, and HCV RNA was extracted using the NucliSENS easyMAG automated nucleic acid extraction system (bioMérieux, Marcy l'Etoile, France), reverse-transcribed, and amplified using sets of newly designed and previously published primers (10) spanning the complete genome of HCV that encodes the structural proteins (C, E1, and E2), non-structural proteins (P7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B), and the 5′- and 3′-untranslated regions (UTRs). The contiguous nucleotide sequences generated by an ABI Prism 3730xl DNA analyzer (Applied Biosystems) were assembled to generate a near-full-length genome of 9,403 bp with an intact open reading frame. Codon alignment was generated with a comprehensive list of reference sequences retrieved from the HCV database (http://www.hcv.lanl.gov/). Maximum likelihood inference was performed using Phylogenetic Analysis Using Parsimony (PAUP) (11), with a Hasegawa-Kishino-Yano nucleotide substitution model and gamma distribution plus discrete gamma categories. In this model, the maximum likelihood phylogenies were heuristically deduced using subtree pruning and regrafting, as well as nearest neighbor interchange algorithms. The statistical robustness and reliability of the branching orders were evaluated by a bootstrap analysis of 1,000 replicates. A neighbor-joining tree was also reconstructed based on the Kimura 2-parameter model implemented in Molecular Evolutionary Genetics Analysis (MEGA) version 5.05 (12).

In the present analysis, both maximum likelihood and neighbor-joining phylogenies showed that the full genome of isolate 10MYKJ032 formed a monophyletic cluster with previously described HCV subtype 6n, namely, isolates KM42 from China (13) and D89/93, TH22, and TH31 from Thailand (10, 14). More remarkably, isolate 10MYKJ032 was located at the root of the 6n lineages with strong statistical support, suggesting its ancestral or parental relationships with other reported HCV 6n genomes. Together with the Thai isolates, phylogenetic inference showed that Southeast Asia may be the plausible geographical origin of HCV subtype 6n. Furthermore, spatial analysis indicated possible migration of HCV 6n lineages from Southeast Asia to the southern regions of China. Such genetic information is important to understand better the evolutionary behaviors of HCV subtype 6n and to trace disease spread in the region.

**Nucleotide sequence accession number.** The genome sequence of 10MYKJ032 has been deposited in GenBank under the accession no. KC191671.

## REFERENCES

1. **Hoofnagle JH.** 2002. Course and outcome of hepatitis C. Hepatology **36**:S21–S29.
2. **World Health Organization.** 17 November 2012, accession date. Hepatitis C fact sheets. World Health Organization, Geneva, Switzerland. http://www.who.int/mediacentre/factsheets/fs164/en/index.html.
3. **Nelson PK, Mathers BM, Cowie B, Hagan H, Des Jarlais D, Horyniak D, Degenhardt L.** 2011. Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. Lancet **378**:571–583.
4. **Simmonds P, Holmes EC, Cha TA, Chan SW, McOmish F, Irvine B, Beall E, Yap PL, Kolberg J, Urdea MS.** 1993. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J. Gen. Virol. **74**:2391–2399.
5. **Murphy DG, Willems B, Deschênes M, Hilzenrat N, Mousseau R, Sabbah S.** 2007. 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. **45**:1102–1112.
6. **McOmish F, Yap PL, Dow BC, Follett EA, Seed C, Keller AJ, Cobain TJ, Krusius T, Kolho E, Naukkarinen R.** 1994. Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. J. Clin. Microbiol. **32**:884–892.
7. **Simmonds P.** 2004. Genetic diversity and evolution of hepatitis C virus—15 years on. J. Gen. Virol. **85**:3173–3188.
8. **Pybus OG, Barnes E, Taggart R, Lemey P, Markov PV, Rasachak B, Syhavong B, Phetsouvanah R, Sheridan I, Humphreys IS, Lu L, Newton PN, Klenerman P.** 2009. Genetic history of hepatitis C virus in East Asia. J. Virol. **83**:1071–1082.
9. **Simmonds P.** 1997. Clinical relevance of hepatitis C virus genotypes. Gut **40**:291–293.
10. **Lu L, Li C, Fu Y, Thaikruea L, Thongswat S, Maneekarn N, Apichart-piyakul C, Hotta H, Okamoto H, Netski D, Pybus OG, Murphy D, Hagedorn CH, Nelson KE.** 2007. Complete genomes for hepatitis C virus subtypes 6f, 6i, 6j and 6m: viral genetic diversity among Thai blood donors and infected spouses. J. Gen. Virol. **88**:1505–1518.
11. **Swofford DL.** 2003. PAUP*: phylogenetic analysis using parsimony (and other methods) 4 beta. Sinauer Associates, Sunderland, MA.
12. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**:2731–2739.
13. **Lu L, Nakano T, Li C, Fu Y, Miller S, Kuiken C, Robertson BH, Hagedorn CH.** 2006. Hepatitis C virus complete genome sequences identified from China representing subtypes 6k and 6n and a novel, as yet unassigned subtype within genotype 6. J. Gen. Virol. **87**:629–634.
14. **Noppornpath S, Poovorawan Y, Lien TX, Smits SL, Osterhaus AD, Haagmans BL.** 2008. Complete genome analysis of hepatitis C virus subtypes 6t and 6u. J. Gen. Virol. **89**:1276–1281.