HEPATITIS C VIRUS PHYLOGENY: A USEFUL CLINICAL TOOL

J. CRISTINA[1*], S. MUKOMOLOV[2], R. COLINA[1,3], O. KALININA[2], L. GARCÍA[1], B. KHAN[4], C. MOGDASY[3], P. KARAYIANNIS[5]

[1]Departamento de Técnicas Nucleares Aplicadas, Centro de Investigaciones Nucleares, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay; [2]Saint Petersburg Pasteur Institute, Saint Petersburg, Russia; [3]Laboratorio de Biología Molecular, Asociación Española Primera de Socorros Mutuos, Montevideo, Uruguay; [4]Division of Human Health, International Atomic Energy Agency, Vienna, Austria; [5]Department of Medicine A, Faculty of Medicine, Imperial College of Science, Technology and Medicine, London, United Kingdom

Received August 19, 2002; accepted September 2, 2002

Summary. – In order to type 45 recent isolates of Hepatitis C virus (HCV) originating from four different geographic regions of the world, we performed phylogenetic analysis of a 192 nucleotides (nts) long sequence from the 5'non-coding region (5'-NCR) of the virus genome and compared them with 55 HCV isolates/strains of known type. The results of this study showed that phylogenetic studies can assign an HCV isolate to the correct type in 100% and to the correct subtype in 98%. A comparison of this method with other methods using commercial kits revealed that it is appropriate for clinical use and is cost effective.

Key words: Hepatitis C virus; genotypes; genetic variability; virus typing; phylogenetic tree analysis

HCV is a major cause of post-transfusion sporadic non-A, non-B hepatitis throughout the world (Manzin *et al.*, 1998). About a half of all patients with acute hepatitis C progress to chronic disease (Farci *et al.*, 1991) and many of them eventually develop hepatocellular carcinoma (Saito *et al.*, 1990). HCVs can be classified into six major genotypes, many of which contain a number of more closely related subtypes (Simmonds *et al.*, 1993).

Several methodologies have been developed for identification of HCV types and subtypes. These include InnoLipa HCV II (Innogenetics, Belgium) (Shobokshi *et al.*, 1999; Sandres *et al.*, 2001), the Murex serotyping 1-6 assay (Murex, UK) (Sandres *et al.*, 2001), Chiron RIBA HCV 3.0 Strip Immunoblot (Chiron Corporation, USA) (Sandres *et al.*, 2001; Fabrizi *et al.*, 2000), and the Trugene Hepatitis C assay (Visible Genetics, USA) (Ansaldi *et al.*, 2001). However, such assay kits are expensive and are beyond the economic means of most developing countries of the world.

*E-mail: cristina@cin1.cin.edu.uy; fax: +5982-5250895.

Abbreviations: HCV = Hepatitis C virus; 5'-NCR = 5'-non-coding region; nt = nucleotide; PCR = polymerase chain reaction

In order to test the capacity of phylogenetic analysis to assign HCV isolates to the correct type and subtype, we analyzed 45 recent HCV isolates originating from Russia, the United Kingdom, Greece and Uruguay. Nucleotide sequences from the 5'-NCR were obtained by polymerase chain reaction (PCR) (Chan *et al.*, 1992) and either manual (Sequenase version 2.0, Amersham BioSciences, UK) or automated sequencing in an ABI Prism 377 version 3.0 automated DNA sequencer. In addition, the genotype and subtype of these isolates was also determined by restriction fragment length polymorphism analysis of the 5'-NCR (Davidson *et al.*, 1995; Okamoto *et al.*, 1992) (data not shown).

The results of these studies are shown in Table 1. To determine the reliability of the phylogenetic analysis in accurately assigning HCV type and subtype, the sequences obtained from the 5'-NCR (nts 121–313) of the genomes of the isolates were aligned with the corresponding ones from 55 strains of known type/subtype reported by others using the Clustal W Program (Higgins *et al.*, 1996). Then, using the MEGA Program Version 2.0 (Kumar *et al.*, 1984), a matrix of distances was created for the Kimura-two parameters model (Felsenstein, 1993) showing that the differences in distances ranged from 0 to 16.30%, with an average distance difference of 3.8% among all isolates. Using this matrix and the neighbour-joining method, phylogenetic trees were created and their reliability was assessed by bootstrap resampling (1000 pseudo-replicas). These methods were implemented with the software from the MEGA2 Program.

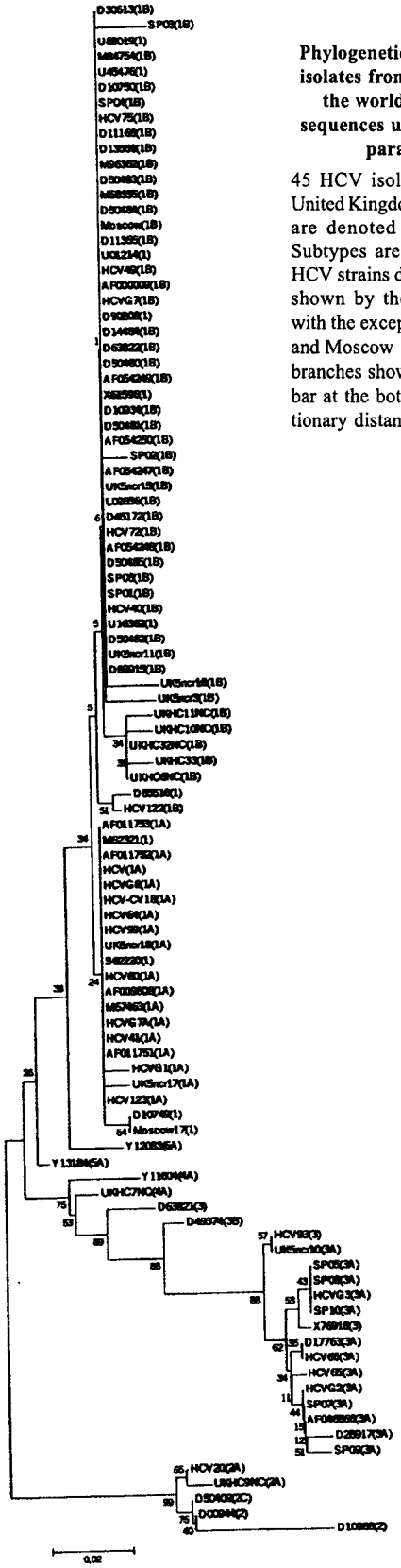

Fig. 1

Phylogenetic analysis of 45 HCV isolates from different regions of the world based on 5'-NCR sequences using the Kimura-two parameter model

45 HCV isolates from Greece, the United Kingdom, Russia and Uruguay are denoted as given in Table 1. Subtypes are given in parentheses. HCV strains described previously are shown by their accession number, with the exception of strains Moscow and Moscow 17. The numbers on the branches show bootstrap values. The bar at the bottom denotes the evolutionary distance.

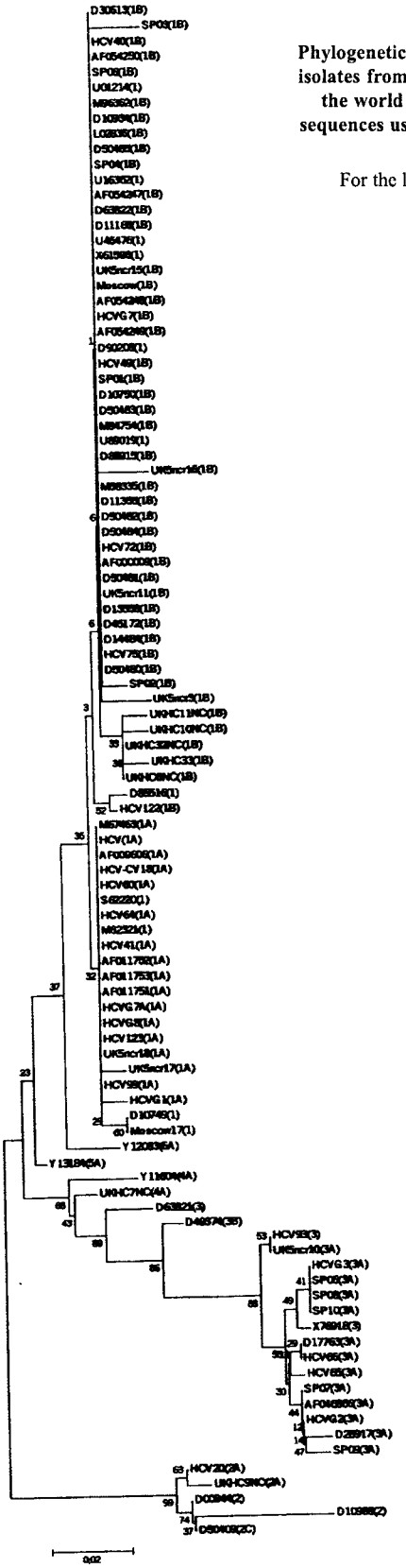

Fig. 2

Phylogenetic analysis of 45 HCV isolates from different regions of the world based on 5'-NCR sequences using the Tamura-Nei model

For the legend see Fig.1.

Table 1. Genotype, designation and origin of the HCV isolates included in this study

Genotype	Name	Origin
1a	HCVG7A, HCV99, HCV, HCV60, HCV64 HCVG8, HCV123, HCV41, HCVG1, HCV-CV18	Montevideo, Uruguay
1b	HCVG7, HCV40,HCV72, HCV75, HCV122	Montevideo, Uruguay
2a	HCV20	Montevideo, Uruguay
3a	HCVG2, HCVG3, HCV65, HCV66,HCV93	Montevideo, Uruguay
1b	SP01, SP02, SP03, SP04, SP06	Saint Petersburg, Russia
	Moscow, Moscow17	Moscow, Russia
3a	SP05, SP07, SP08, SP09, SP10	Saint Petersburg, Russia
1a	UK5ncr16, UK5ncr17, UK5ncr18	United Kingdom
1b	UKncr15, UK5ncr11, UK5ncr3	United Kingdom
3a	UK5ncr1	United Kingdom
1b	UKHC6NC, UKHC10NC, UKHC11NC, UKHC33NC, UKHC32NC	Greece
2a	UKHC9NC	Greece
4a	UKHC7NC	Greece

The results of these studies are shown in Fig. 1. The majority of the isolates included in this study, which are known to belong to type 1, clustered together. The isolates belonging to the other five HCV types clustered separately. Inside of the main cluster of type 1 isolates, two branches could be observed, one representing subtype b (Fig. 1, top and Table 1), and the other representing subtype a (Fig. 1, middle). Thus, all type 1 isolates (64% of the total) could be clearly assigned to a specific subtype. In the case of type 2 isolates a close genetic relationship was observed among all of them, these being more distantly related with the rest of HCV types (Fig. 1, bottom). Two type 2A isolates from this study (HCV20 and UKHC9NC) were positioned on the same branch of type 2 isolates and also show a closer genetic relationship among themselves in comparison with type 2C and other type 2 isolates.

A cluster of type 3A isolates with a high bootstrap value could also be clearly seen (Fig. 1, bottom). Nevertheless, the genetic distances among this group of isolates and other type 3 isolates were higher in comparison with those among different subtypes within a given HCV type. For this reason, using this approach, it is not clear whether the strain D63821 (previously described) is type 3b or belongs to another subtype. The same is true for the strain UKHC7NC, which is genetically close to the prototype 4a strain and has been previously typed by us as 4a by another approach. The current phylogenetic analysis does not permit a clear assignment to type 4a. Probably more type 4a strains need be included to establish a definitive picture. Types 5 and 6 are represented as separate clusters with only the prototype strains, since no such isolates belonging to these genotypes were obtained from our patients. In spite of the short length of the amplicon (192 nts) and relatively few nucleotide differences between subtypes (1a and 1b, as well as subtypes of type 2), subtype assignment was accurate in 98% of cases.

To confirm these findings phylogenetic analyses were done using the Tamura-Nei matrix of distances (Tamura and Nei, 1993) and the neighbor-joining method. In this way a phylogenetic tree was generated, the reliability of which was also tested by bootstrap resampling (1000 pseudo-replicas). The results were the same as those obtained with the Kimura-two parameter model (Fig. 2). As the need for more HCV genotyping increases, rapid and reliable methods are needed for accurate assignment of a large number of isolates. Phylogenetic analysis has the advantage that statistical as well as phylogenetic measures can be used for a more appropriate description of genetic relationships among different HCV isolates/strains, as well as for determination of their evolutionary and epidemiological history. It is cost effective, and therefore within reach of laboratories in the developing world, as the available kits are currently too expensive. Using this approach, all 45 HCV isolates investigated in this study could be assigned to a correct genotype, whilst 98% of them were correctly subtyped.

In a recent study, comparison of the results obtained by sequencing the NS5b gene with those obtained by the Trugene and InnoLipa assays showed concordance for HCV types in 100% of cases, whilst the ability of these methods to discriminate at the subtype level was 76% and 74% for the Trugene and Innolipa assays, respectively (Halfon *et al.*, 2001). Taking these results into account, the phylogenetic method employed here turned out to be sensitive, rapid, low cost and suitable for accurate large-scale genotyping of HCV.

Acknowledgement. The authors acknowledge support from the International Atomic Energy Agency through Coordinated Research Project E1.50.18 and research contracts Nos. 10868, 10867, 10978.

References

Ansaldi F, Torre F, Bruzzone BM, Picciotto A, Crovari P, Icardi G (2001): Evaluation of a new hepatitis C virus sequencing assay as a routine method for genotyping. *J. Med. Virol.* **63**, 17–21.

Chan SW, McOmish F, Holmes EC, Dow B, Peutherer JF, Follett E, Yap PL, Simmonds P (1992): Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. *J. Gen. Virol.* **73**, 1131–1141.

Davidson F, Simmonds P, Ferguson J, Jarvis L, Dow B, Follett E, Seed C, Krusius T, Lin C, Medgyesu G, Kiyokawa H, Olim G, Duraisamy G, Cuypers T, Seed A, Ten D, Conradie J, Kew M, Lin M, Nuchaprayoon C, Ndimbe O, Yap PL (1995): Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding regions. *J. Gen. Virol.* **76**, 1197–1204.

Fabrizi F, Martin P, Quan S, Dixit V, Brezina M, Conrad A, Polito A, Gitnick G (2000): Serotyping strip immunoblot assay for assessing hepatitis C virus strains in dialysis patients. *Am. J. Kidney Dis.* **35**, 832–838.

Farci P, Alter AJ, Wong D, Miller RH, Shin JW, Jett B, Purcell RH (1991): A long-term study of hepatitis C virus replication in non-A, non-B hepatitis. *N. Engl. J. Med.* **325**, 98–104.

Felsenstein J (1993): Phylogeny inference package, version 3.5. Department of Genetics, University of Washington, Seattle, USA.

Halfon P, Trimoulet P, Bourliere M, Khiri H, de Ledinghen V, Couzigou P, Feryn JM, Alcaraz P, Renou C, Fleury HJ, Ouzan D (2001): Hepatitis C virus genotyping based on 5' noncoding sequence analysis (Trugene). *J. Clin. Microbiol.* **39**, 1771–1773.

Higgins DG, Thompson JD, Gibson TJ (1996): Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266**, 383–402.

Kumar S, Tamura K, Nei M (1994): MEGA: molecular evolutionary genetics analysis for microcomputers. *Comput. Appl. Biosci.* **10**, 189–191.

Manzin A, Solforosi L, Petrelli E, Macarri G, Grazia T, Piazza M, Clementi M (1998): Evolution of hypervariable region 1 of hepatitis C virus in primary infection. *J. Virol.* **72**, 6271–6276.

Okamoto H, Sujiyama S, Okada K, Kurai K, Akahane Y, Sugai Y, Tanaka T, Sato K, Tsuda F, Miyakawa Y, Mayumi M (1992): Typing hepatitis C virus by polymerase chain reaction with type-specific primers: application to clinical surveys and tracing infection sources. *J. Gen. Virol.* **73**, 673–679.

Saito I, Miyamura T, Ohbayashi A, Harada H, Katayama T, Kikuchi S, Watanabe Y, Koi S, Onji M, Ohta Y, Choo Q, Houghton M, Kuo G (1990): Hepatitis C virus infection is associated with development of hepatocellular carcinoma. *Proc. Natl. Acad. Sci. USA* **87**, 6547–6549.

Sandres K, Dubois M, Pasquier C, Puel J, Izopet J (2001): Determination of HCVgenotype using two antibody assays and genome typing. *Eur. J. Clin. Microbiol. Infect. Dis.* **20**, 666–669.

Shobokshi O, Serebour F, Skakni L, Al-Saffy Y, Ahdal M (1999): Hepatitis C genotypes and subtypes in Saudi Arabia. *J. Med. Virol.* **58**, 44–48.

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.

Tamura K, Nei M (1993): Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526.